

Calibration of a Four-Factor Hybrid Local-Stochastic Volatility Model with a New Control Variate Particle Method

ANDREI COZMA* MATTHIEU MARIAPRAGASSAM*[†] ✉
CHRISTOPH REISINGER*[†]

Abstract

We propose a novel and generic calibration technique for four-factor foreign-exchange hybrid local-stochastic volatility models (LSV) with stochastic short rates. We build upon the particle method introduced by Guyon and Labordère [Nonlinear Option Pricing, Chapter 11, Chapman and Hall, 2013] and combine it with new variance reduction techniques in order to accelerate convergence. We use control variates derived from a calibrated pure local volatility model, a two-factor Heston-type LSV model (both with deterministic rates), and the stochastic (CIR) short rates. Our numerical experiments show that because of the dramatic variance reduction we are able to calibrate the four-factor model at almost no extra computational cost when the corresponding calibrated two-factor model is at our disposal. The method can be applied to a large class of hybrid LSV models and is not restricted to our particular choice of the diffusion. The calibration procedure is performed on real-world market data for the EUR-USD currency pair.

1 Introduction

Efficient pricing and hedging of exotic derivatives requires a model which is rich enough to re-price accurately a range of liquidly traded market products. The case of calibration to vanilla options is now widely documented and has been considered extensively in the literature since the work of Dupire [18] in the context of local volatility models (see also Gyöngy’s formula [27]). Nowadays, the exact re-pricing of call options is a must-have standard, and *Local-Stochastic Volatility* (LSV) models are the *state-of-the-art* in many financial institutions because of their superior dynamic properties over pure local volatility models. Various sophisticated calibration techniques are in use in the financial industry, for example, based on the work of Guyon and Labordère [25] as well as Ren, Madan and Qian [38]. According to Ren et al. [38], Tian et al. [44], Stoep et al. [46] and Guyon and Labordère [25], the general LSV model allows for a better calibration to European options and improves the pricing and risk-management performance when compared to pure local volatility or pure stochastic volatility models. We focus on the Heston-type LSV model because the Cox–Ingersoll–Ross (CIR) process for the variance is widely used in the industry due to its desirable properties, such as mean-reversion and non-negativity, and since semi-analytic formulae are available for calls and puts under Heston’s model and can help calibrate the parameters easily. The local volatility component allows a perfect calibration to the market prices of vanilla options. At the same time, the stochastic volatility component already provides built-in smiles and skews

*MATHEMATICAL INSTITUTE, UNIVERSITY OF OXFORD, OX2 6GG, UNITED KINGDOM

[†]OXFORD-MAN INSTITUTE OF QUANTITATIVE FINANCE, UNIVERSITY OF OXFORD, OX2 6ED, UNITED KINGDOM
{ANDREI.COZMA, MATTHIEU.MARIAPRAGASSAM, CHRISTOPH.REISINGER}@MATHS.OX.AC.UK

THE FIRST AUTHOR GRATEFULLY ACKNOWLEDGES FINANCIAL SUPPORT FROM THE EPSRC. THE SECOND AUTHOR GRATEFULLY ACKNOWLEDGES FINANCIAL SUPPORT FROM THE OXFORD–MAN INSTITUTE AND BNP PARIBAS LONDON.

which give a rough fit, so that a leverage function relatively close to one suffices for a perfect calibration.

In order to improve the pricing and hedging of foreign exchange (FX) options, we introduce stochastic domestic and foreign short interest rates into the model. Assuming constant interest rates is appealing due to its simplicity and does not lead to serious mispricing of options with short maturities. However, empirical results [48] have confirmed that the constant interest rate assumption is inappropriate for long-dated FX products, and the effect of interest rate volatility can be as relevant as that of the FX rate volatility for longer maturities. Extensive research has been carried out in the area of option pricing with stochastic volatility and interest rates in the past few years. Van Haastrecht et al. [48] extended the model of Schöbel and Zhu [40] to currency derivatives by including stochastic interest rates, a model that benefits from analytical tractability even in a full correlation setting due to the processes being Gaussian. On the other hand, Ahlip and Rutkowski [2], Grzelak and Oosterlee [24] and Van Haastrecht and Pelsser [49] examined the Heston–CIR/Vasicek hybrid models and concluded that they give rise to non-affine models even under a partial correlation structure of the driving Brownian motions.

In the foreign exchange market, the inclusion of stochastic interest rates results in a 4-factor model associated with a complex calibration routine. A few papers discuss this problem under simpler settings. Deelstra [16] and Clark [10] mainly consider 3-factor hybrid local volatility models and focus on the theoretical rather than the practical aspects of the calibration, whereas Stoep, Grzelak and Oosterlee [47] consider an application to a 2-factor hybrid local volatility. In [25], Guyon and Labordère discuss an application of Monte Carlo-based calibration methods to a 3-factor LSV equity model with stochastic domestic rate and discrete dividends.

The model of Cox et al. [13] is very popular when modeling short rates because the (square-root) CIR process admits a unique strong solution, is mean-reverting and analytically tractable. Cox et al. found the conditional distribution to be non-central chi-squared, on the basis of which Broadie and Kaya [7] proposed an efficient exact simulation scheme for the (square-root) process using acceptance-rejection sampling. As of late, the non-negativity of the CIR process is considered to be less desirable when modeling short rates. On one hand, central banks have significantly reduced the interest rates since the 2008 financial crisis and it is now commonly accepted that interest rates need not be positive. On the other hand, if interest rates dropped too far below zero, then large amounts of money would be withdrawn from banks and government bonds, putting a severe squeeze on deposits. Hence, we model the domestic and foreign short rates using the shifted CIR (CIR++) process of Brigo and Mercurio [5]. The CIR++ model allows the short rates to become negative and can fit any observed term structure exactly while preserving the analytical tractability of the original model for bonds, caps, swaptions and other basic interest rate products.

In this work, we consider the 4-factor hybrid LSV model defined below in (2.1), which is a Heston-type LSV model with two shifted CIR short-rate processes that accounts for negative rates and provides alternative modeling of rates in lieu of the more classical Hull–White process. We extend the calibration condition for the leverage function given in [25] for a 3-factor LV–2Hull–White model and in [16] for a 4-factor Heston-type LSV–2Hull–White model to the 4-factor LSV–2CIR++ model, and give a novel and rigorous derivation.

We propose a new calibration approach that builds on the particle method of [25], and combine it with a novel and efficient variance reduction technique that takes advantage of PDE calibration to increase its stability and accuracy while keeping the dimensionality under control with Monte Carlo.

The proposed variance reduction techniques allow to evaluate each expectation with a small amount of particles/paths (we find that 4000 – 20000 particles are sufficient in practice). Our numerical experiments suggest that this method almost recovers the calibration speed from the corresponding 2-factor LSV model with deterministic rates defined in (2.2).

As a side result, we explain how to effectively deal with the broken Feller condition for the Heston-type LSV Kolmogorov forward *Partial Differential Equation* (PDE) and numerically solve the PDE using the finite element method with a *Backward Differentiation Formula* (BDF) time-stepping scheme and appropriate non-Dirichlet boundary condition. To the best of our knowledge, this represents a new approach that complements the literature on the use of ADI schemes [10, 38, 52] to

handle the PDE calibration of an LSV model with deterministic rates. Moreover, a main advantage of the finite element method compared to ADI schemes is that the Dirac delta initial condition can be handled naturally in the weak formulation. This framework yields an accurate calibration of an LSV with deterministic rates for a broad set of market data. The idea we propose in this article is to use the 2-factor LSV model calibrated via the PDE as a “control variate” (generalising this concept from the standard setting to conditional expectations) for a 4-factor LSV model with stochastic rates calibrated via a Monte Carlo particle method as outlined above.

Furthermore, we rely on the forward Dupire PDE to compute the second order derivative of the call price with respect to the strike, which appears in the calibration formula (similar to the Dupire formula), instead of extracting this quantity from the market. By doing so, we prevent instabilities and get a more accurate estimate of the leverage function, which is only possible through a proper calibration of a pure local volatility model that we will also discuss.

Finally, we believe that including stochastic rates is of paramount importance to price some specific exotic derivatives such as the Accumulator, the *Power Reverse Dual-Currency note* (PRDC) or the *AutoCallable Barrier Dual-Currency Note* (ABDC). In particular, in [14] we consider the pricing problem for the last contract and show the impact of stochastic rates on the net present value. Both the AutoCallable and the Accumulator can embed a barrier feature that will benefit from a local-stochastic volatility component. Indeed, pure stochastic volatility models tend to underestimate the knock-out probability while pure local volatility models over-estimate it. Therefore, a combination of both local and stochastic volatility provides the best trade-off for which a mixing factor can eventually be included [10]. Moreover, stochastic rates become necessary for any hybrid product that embeds the rates explicitly. One example is the spread option between the performance of the FX rate and the Libor rate.

The remainder of this paper is organised as follows. In Section 3, we provide a necessary and sufficient condition for a perfect calibration to vanilla quotes and prove this rigorously emphasising the use of local times and possible moment explosions. In Section 4, we present our approach to the calibration problem and describe the process step-by-step. First, we calibrate a pure local volatility model (2.3) with the forward Dupire PDE. Second, we propose to calibrate a LSV model with deterministic rates using the finite element method. Furthermore, we employ a change-of-variable technique to tackle the broken Feller condition efficiently. In Section 5, we combine the calibrated 2-factor LSV model (2.2) and variance reduction techniques to calibrate the 4-factor LSV model (2.1) with stochastic rates via the particle method. We present numerical results and show that a low number of particles suffice to provide a very good fit to market quotes, which suggests that the method is computationally efficient. Section 6 concludes with a brief discussion.

2 Model definition

We consider a domestic and a foreign market with risk-free short rates r^d and r^f . We denote by D^d and D^f the domestic and foreign discount factors associated with their respective money market accounts. Throughout, T and t are fixed positive times such that $t < T$. We assume the existence of a filtered probability space $(\mathcal{X}, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{Q}^d)$ with a domestic risk-neutral measure \mathbb{Q}^d , under which the price process of a risky asset follows a system of SDEs

$$\begin{cases} \frac{dS_t}{S_t} = (r_t^d - r_t^f) dt + \alpha(S_t, t) \sqrt{V_t} dW_t \\ r_t^d = g_t^d + h^d(t) \\ r_t^f = g_t^f + h^f(t) \\ dg_t^d = \kappa_d (\theta_d - g_t^d) dt + \xi_d \sqrt{g_t^d} dW_t^{g^d} \\ dg_t^f = \left(\kappa_f (\theta_f - g_t^f) - \rho_{Sf} \xi_f \sqrt{g_t^f} \alpha(S_t, t) \sqrt{V_t} \right) dt + \xi_f \sqrt{g_t^f} dW_t^{g^f} \\ dV_t = \kappa (\theta - V_t) dt + \xi \sqrt{V_t} dW_t^V \end{cases} \quad (2.1)$$

with the correlation structure

$$d\langle W_t, W_t^V \rangle = \rho dt, \quad d\langle W_t, W_t^{g^d} \rangle = \rho_{Sd} dt, \quad d\langle W_t, W_t^{g^f} \rangle = \rho_{Sf} dt,$$

with $\rho, \rho_{Sd}, \rho_{Sf} \in (-1, 1)$, the other correlations being zero, and for given functions $\alpha : \mathbb{R}^+ \times [0, T] \rightarrow \mathbb{R}^+$, $h^{d/f} : [0, T] \rightarrow \mathbb{R}$, and non-negative numbers $\kappa, \theta, \xi, \kappa_d, \theta_d, \xi_d, \kappa_f, \theta_f, \xi_f$, as well as initial values S_0, g_0^d, g_0^f, V_0 .

The domestic and foreign discount factors are then given by

$$D_t^d = e^{-\int_0^t r_u^d du}, \quad D_t^f = e^{-\int_0^t r_u^f du},$$

respectively.

We also define two simpler models that we will refer to in the remainder of the article. First, we can write the related 2-factor Heston-type LSV model with deterministic rates as

$$\begin{cases} \frac{dS_t^{2D}}{S_t^{2D}} = (\bar{r}^d(t) - \bar{r}^f(t)) dt + \alpha^{2D}(S_t^{2D}, t) \sqrt{V_t} dW_t, & S_0^{2D} = S_0, \\ dV_t^{2D} = \kappa(\theta - V_t^{2D}) dt + \xi \sqrt{V_t^{2D}} dW_t^V \end{cases} \quad (2.2)$$

for a given function $\alpha^{2D} : \mathbb{R}^+ \times [0, T] \rightarrow \mathbb{R}^+$, and the pure *Local volatility* (LV) model as

$$\frac{dS_t^{LV}}{S_t^{LV}} = (\bar{r}^d(t) - \bar{r}^f(t)) dt + \sigma_{LV}(S_t^{LV}, t) dW_t, \quad S_0^{LV} = S_0, \quad (2.3)$$

where $\bar{r}^d(t) = -\frac{\partial \ln P^d(0, t)}{\partial t}$ and $\bar{r}^f(t) = -\frac{\partial \ln P^f(0, t)}{\partial t}$, with $P^{d/f}(0, T)$ the market zero coupon bond prices for the domestic and foreign money market accounts, respectively, and a given function $\sigma_{LV} : \mathbb{R}^+ \times [0, T] \rightarrow \mathbb{R}^+$. Note that while the volatility is ‘local’, i.e., a function of spot FX and time, the short rates are assumed to be a function of time only. We also note for future reference that under the pure LV model (2.3) call prices C_{LV} satisfy the forward Dupire PDE

$$\frac{\partial C_{LV}}{\partial T} - \frac{1}{2} \sigma_{LV}(K, T)^2 K^2 \frac{\partial^2 C_{LV}}{\partial K^2} + K(\bar{r}^d(T) - \bar{r}^f(T)) \frac{\partial C_{LV}}{\partial K} + \bar{r}^f(T) C_{LV} = 0. \quad (2.4)$$

The purpose of this paper is to calibrate $h^d, h^f, \kappa_d, \kappa_f, \theta_d, \theta_f, \xi_d, \xi_f, \kappa, \theta, \xi, \rho$ and especially α in (2.1). We will use calibration of (2.2) and (2.3) as “stepping stones”. In order to illustrate the full calibration process from start to finish and to make the structure clear to the reader, we provide a flowchart of the calibration in Figure 2.1.

3 A necessary and sufficient condition for exact calibration

In this section, we derive the main formula that links market call prices, via the Dupire local volatility, to call prices under model (2.1).

The spot S_T is associated with the currency pair ccy1ccy2 (following the notations in [10]) and denotes the amount of units of ccy2 (domestic currency) needed to buy one unit of ccy1 (foreign currency) at time T .

Let the discounted call option price under model (2.1) for a notional of one unit of ccy1 be

$$C(K, T) = \mathbb{E}^{\mathbb{Q}^d} \left[D_T^d (S_T - K)^+ \right].$$

The owner of a long call contract, for a notional amount N_{ccy1} , has the right to buy N_{ccy1} in ccy1 in exchange for an amount $N_{ccy2} = KN_{ccy1}$ in ccy2 at expiry. The payout is $(N_{ccy1}S_T - N_{ccy2})^+$ and its discounted price is $N_{ccy1}C(K, T)$. The market standard is to quote call options in ccy2.

We can now state the following result, which we will prove in the remainder of this section.

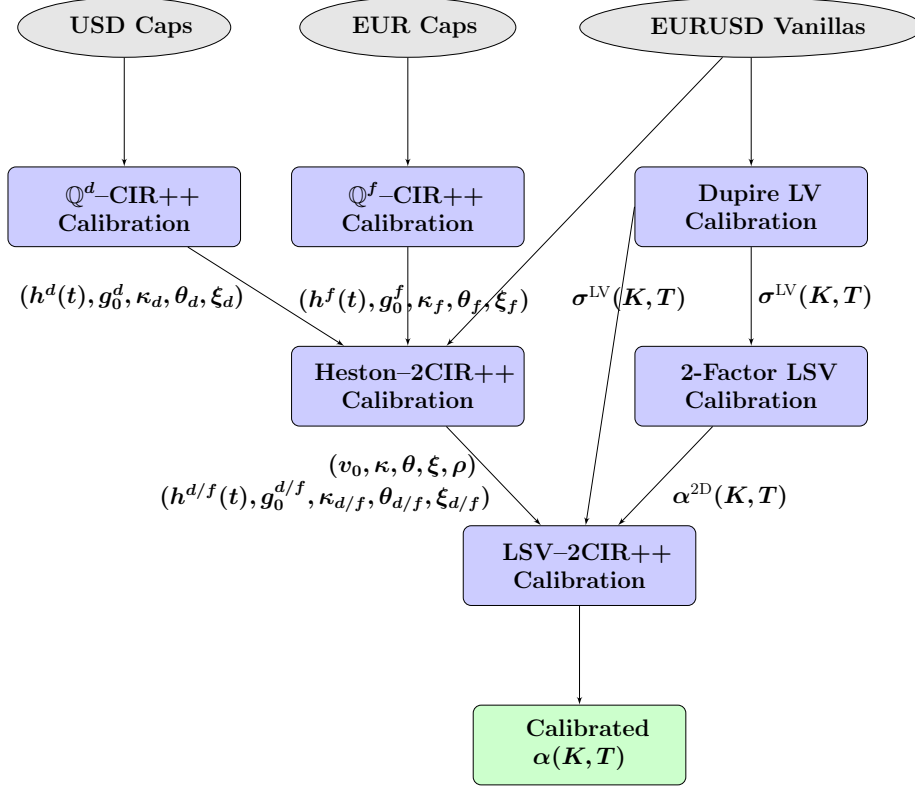


Figure 2.1: Full calibration routine flowchart

Theorem 1. Assume that α is uniformly bounded by α_{max} , $h^{d,f}$ are uniformly bounded and that both the marginal density $\phi(T, \cdot)$ of S_T in (2.1) and $\mathbb{E}^{\mathbb{Q}^d} [D_T^d V_T | S_T = \cdot]$ are continuous.

Then, the call price $C(K, T)$ under model (2.1) matches the price C^{LV} under the local volatility model (2.3) for any strike $K > 0$ and maturity $T < T^*$ if and only if, for all $K, T > 0$,

$$\alpha^2(K, T) = \frac{\mathbb{E}^{\mathbb{Q}^d} [D_T^d | S_T = K]}{\mathbb{E}^{\mathbb{Q}^d} [D_T^d V_T | S_T = K]} \left(\sigma_{LV}^2(K, T) + q(K, T) \right), \quad (3.1)$$

where σ_{LV} is a local volatility as in (2.3),

$$q(K, T) = \frac{\mathbb{E}^{\mathbb{Q}^d} [Q_T]}{\frac{1}{2} K^2 \frac{\partial^2 C_{LV}}{\partial K^2}}, \quad (3.2)$$

$$Q_T = D_T^d \left(r_T^f - \bar{r}^f(T) \right) (S_T - K)^+ - K D_T^d \mathbf{1}_{S_T \geq K} \left[\left(r_T^d - \bar{r}^d(T) \right) - \left(r_T^f - \bar{r}^f(T) \right) \right],$$

and, denoting $\varphi = 2 + \sqrt{2}$ and $\zeta = \xi \alpha_{max}$,

$$\begin{cases} T^* = \frac{2}{\sqrt{\varphi^2 \zeta^2 - \kappa^2}} \left[\frac{\pi}{2} + \arctan \left(\frac{\kappa}{\sqrt{\varphi^2 \zeta^2 - \kappa^2}} \right) \right], & \text{if } \kappa < \varphi \zeta, \\ T^* = \infty, & \text{if } \kappa \geq \varphi \zeta. \end{cases} \quad (3.3)$$

Remark. The ratio on the right-hand side of (3.1) accounts for the stochastic volatility, such that if there is no stochastic volatility (i.e. $V_T = 1$), we recover the formula in [10]. The term q in (3.2) accounts for the stochasticity of rates and, if rates are deterministic, $q = 0$ and we recover the formula derived in [38, 25]

$$\alpha^{2D}(K, T) = \frac{\sigma_{LV}(K, T)}{\sqrt{\mathbb{E}^{\mathbb{Q}^d} [V_T^{2D} | S_T^{2D} = K]}}. \quad (3.4)$$

At time $T = 0$, since by construction we have $r_0^d - \bar{r}^d(0) = r_0^f - \bar{r}^f(0) = 0$, the leverage function is $\alpha(K, 0) = \frac{\sigma_{LV}(K, 0)}{\sqrt{v_0}}$.

An alternative form of (3.1) can be obtained by subtracting $\alpha(K, T) \frac{\mathbb{E}^{\mathbb{Q}^d}[D_T^d V_T | S_T = K]}{\mathbb{E}^{\mathbb{Q}^d}[D_T^d | S_T = K]} \times$ (2.4) from $\sigma_{LV}^2(K, T) \times$ (3.6), namely

$$\begin{aligned} \alpha^2(K, T) &= \sigma_{LV}^2(K, T) \frac{\mathbb{E}^{\mathbb{Q}^d}[D_T^d | S_T = K]}{\mathbb{E}^{\mathbb{Q}^d}[D_T^d V_T | S_T = K]} \\ &\times \left(\frac{\frac{\partial C}{\partial T} + \mathbb{E}^{\mathbb{Q}^d}[D_T^d r_T^f (S_T - K)^+] - \mathbb{E}^{\mathbb{Q}^d}[D_T^d \mathbf{1}_{S_T \geq K} K (r_T^d - r_T^f)]}{\frac{\partial C}{\partial T} + \bar{r}^f(T) C_{LV} + K (\bar{r}^d(T) - \bar{r}^f(T)) \frac{\partial C_{LV}}{\partial K}} \right). \end{aligned} \quad (3.5)$$

If we divide the denominator by σ_{LV}^2 and re-arrange, we recover the formula presented in [16] for an LSV-2Hull-White process, namely

$$\alpha^2(K, T) \frac{\mathbb{E}^{\mathbb{Q}^d}[D_T^d V_T | S_T = K]}{\mathbb{E}^{\mathbb{Q}^d}[D_T^d | S_T = K]} = \frac{\frac{\partial C}{\partial T} - \mathbb{E}^{\mathbb{Q}^d}[D_T^d (K r_T^d - r_T^f S_T) \mathbf{1}_{S_T > K}]}{\frac{1}{2} K^2 \frac{\partial^2 C_{LV}}{\partial K^2}}.$$

However, we believe practitioners are more familiar with (3.1) as derived in [10] for a 3-factor LV-2Hull-White process, which is why we use the main formulation (3.1) in the remainder of the article.

Theorem 1 generalises the formula presented in [25] to stochastic foreign rates. Here, we consider an LSV-2CIR++ model and derive our result in a rigorous fashion, emphasising the importance of the local time process. Furthermore, in Lemma 3 we provide a sufficient condition for the process

$$\int_0^t \mathbf{1}_{S_u \geq K} D_u^d S_u \alpha(S_u, u) \sqrt{V_u} dW_u$$

to be a true martingale up to T^* , which is an important step in the proof of Theorem 1 that is often omitted in the literature. On the one hand, T^* is a lower bound for the explosion time of the second moment of the discounted spot process $D_t^d S_t$. On the other hand, [4] show that the moment explodes in finite time for the Heston model, a property that is inherited by our Heston-type LSV-2CIR++ model (2.1) as well as the Heston-type LSV-2Hull-White in [16]. Therefore, the formula may not hold for some values of the model parameters and for large maturities T . However, in practice, T^* is very large. For instance, following our calibration routine in Section 5 we obtain

$$\kappa = 1.4124, \xi = 0.2988, \alpha_{max} = 1.40,$$

such that

$$T^* = 28.6.$$

An issue may arise for large values of ξ . Indeed, a numerical experiment in [25], raises a question on the existence of a calibrated 2-factor LSV model when $\xi \approx 350\%$ (high values of the vol-of-vol are used to match forward smiles). In this particular case and with the other model parameters kept the same, we find $T^* = 0.20$, which indicates that moment explosions may occur sooner.

As an aside, we note that the continuity of the joint density $\psi(T, \cdot, \cdot, \cdot)$ of (S_T, V_T, D_T^d) in (2.1), is sufficient for $\mathbb{E}^{\mathbb{Q}^d}[D_T^d V_T | S_T = \cdot]$ to be continuous in the assumptions of Theorem 1.

We first state and prove two auxiliary lemmas necessary for the derivation of Theorem 1.

Lemma 2. *Given a filtered probability space $(\mathcal{X}, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{Q}^d)$, let W be a standard Brownian motion and μ, Y two \mathcal{F}_t -adapted processes with Y continuous and of finite second moment and*

$$\int_0^t (|\mu_u| + Y_u^2) du < \infty.$$

Consider a continuous Itô process X given by

$$X_t = X_0 + \int_0^t \mu_u du + \int_0^t Y_u dW_u$$

whose marginal density function $\phi(t, \cdot)$ and $\mathbb{E}^{\mathbb{Q}^d} [D_t Y_t^2 | X_t = \cdot]$ are assumed to be continuous. Further denote by $(l_t^a)_{t \geq 0}$ the local time of X at level a . Then, for any continuous, integrable and positive \mathcal{F}_t -adapted semi-martingale D and any real number a ,

$$\mathbb{E}^{\mathbb{Q}^d} \left[\int_0^t D_u dl_u^a \right] = \int_0^t \left(\mathbb{E}^{\mathbb{Q}^d} [D_u Y_u^2 | X_u = a] \phi(a, u) \right) du.$$

Proof. See Appendix A.1.1. □

Lemma 2 provides a link between the local time of a process and its density function. This plays a key role in a financial setting, where the marginal density of the spot is the second order derivative of the call price with respect to the strike, up to a discount factor.

Lemma 3. *Given the set-up of Theorem 1,*

$$M_t = \int_0^t \mathbf{1}_{S_u \geq K} D_u^d \alpha(S_u, u) S_u \sqrt{V_u} dW_u$$

is a true martingale up to T^ given by (3.3).*

Proof. See Appendix A.1.2. □

Combining Lemmas 2 and 3, we can derive the following proposition.

Proposition 4. *Under the assumptions of Theorem 1, and for $T^* > 0$ given by (3.3), the call price $C(K, T)$ under model (2.1) satisfies the forward equation*

$$\begin{aligned} \frac{\partial C(K, T)}{\partial T} - \frac{1}{2} \alpha(K, T)^2 K^2 \frac{\mathbb{E}^{\mathbb{Q}^d} [D_T^d V_T | S_T = K]}{\mathbb{E}^{\mathbb{Q}^d} [D_T^d | S_T = K]} \frac{\partial^2 C(K, T)}{\partial K^2} \\ + \mathbb{E}^{\mathbb{Q}^d} [D_T^d r_T^f (S_T - K)^+] - \mathbb{E}^{\mathbb{Q}^d} [D_T^d \mathbf{1}_{S_T \geq K} K (r_T^d - r_T^f)] = 0 \end{aligned} \quad (3.6)$$

for any strike $K > 0$ and maturity $T < T^$.*

Proof. Let $K \in \mathbb{R}^+$, $0 < t \leq T < T^*$ and $H_t = (S_t - K)^+$. The Trotter-Meyer theorem [39] gives

$$(S_t - K)^+ - (S_0 - K)^+ = \int_0^t \mathbf{1}_{S_u \geq K} dS_u + \frac{1}{2} l_t^K,$$

which we can write in differential form as

$$dH_t = \mathbf{1}_{S_t \geq K} S_t (r_t^d - r_t^f) dt + \frac{1}{2} dl_t^K + \mathbf{1}_{S_t \geq K} \alpha(S_t, t) S_t \sqrt{V_t} dW_t.$$

Also,

$$\begin{aligned} d(D_t^d H_t) &= D_t^d \left[-r_t^d H_t + \mathbf{1}_{S_t \geq K} S_t (r_t^d - r_t^f) \right] dt + \frac{1}{2} D_t^d dl_t^K \\ &+ \mathbf{1}_{S_t \geq K} D_t^d \alpha(S_t, t) S_t \sqrt{V_t} dW_t. \end{aligned} \quad (3.7)$$

Hence, by applying Lemma 2 with $X_t = S_t$, $D_t = D_t^d$ and $Y_t = \alpha(S_t, t) S_t \sqrt{V_t}$, we can write

$$\mathbb{E}^{\mathbb{Q}^d} \left[\int_0^t D_u^d dl_u^K \right] = \int_0^t \left(\alpha(K, u)^2 K^2 \mathbb{E}^{\mathbb{Q}^d} [D_u^d V_u | S_u = K] \phi(K, u) \right) du, \quad (3.8)$$

where ϕ is the marginal density function of S at time t . Furthermore, one can define $\bar{\phi}_n$ as

$$\bar{\phi}_n(K, u) = \mathbb{E}^{\mathbb{Q}^d} [D_u^d \delta_n^K(S_u)] = \int_0^\infty \delta_n^K(x) \mathbb{E}^{\mathbb{Q}^d} [D_u^d | S_u = x] \phi(x, u) dx,$$

with δ_n^K defined as in (A.1) by

$$\delta_n^K(x) = \begin{cases} 0, & |x - K| > \frac{1}{n}, \\ \frac{n}{2}, & |x - K| \leq \frac{1}{n}, \end{cases}$$

and by a similar reasoning to that Lemma 2 we get

$$\lim_{n \rightarrow \infty} \bar{\phi}_n = \mathbb{E}^{\mathbb{Q}^d} [D_u^d | S_u = K] \phi(K, u).$$

And since

$$\frac{\partial^2 C(K, u)}{\partial K^2} = \lim_{n \rightarrow \infty} \mathbb{E}^{\mathbb{Q}^d} [D_u^d \delta_n^K(S_u)],$$

we write

$$\frac{\partial^2 C(K, u)}{\partial K^2} = \mathbb{E}^{\mathbb{Q}^d} [D_u^d | S_u = K] \phi(K, u). \quad (3.9)$$

Combining (3.8) and (3.9) allows to write

$$\mathbb{E}^{\mathbb{Q}^d} \left[\int_0^t D_u^d dL_u^K \right] = \int_0^t \alpha(K, u)^2 K^2 \frac{\mathbb{E}^{\mathbb{Q}^d} [D_u^d V_u | S_u = K]}{\mathbb{E}^{\mathbb{Q}^d} [D_u^d | S_u = K]} \frac{\partial^2 C(K, u)}{\partial K^2} du.$$

Hence, integrating (3.7),

$$\begin{aligned} C(K, t) = \mathbb{E}^{\mathbb{Q}^d} [D_t^d H_t] &= \int_0^t \left(-\mathbb{E}^{\mathbb{Q}^d} [D_u^d r_u^d (S_u - K)^+] + \mathbb{E}^{\mathbb{Q}^d} [D_u^d \mathbf{1}_{S_u \geq K} S_u (r_u^d - r_u^f)] \right) du \\ &+ \frac{1}{2} \int_0^t \alpha(K, u)^2 K^2 \frac{\mathbb{E}^{\mathbb{Q}^d} [D_u^d V_u | S_u = K]}{\mathbb{E}^{\mathbb{Q}^d} [D_u^d | S_u = K]} \frac{\partial^2 C(K, u)}{\partial K^2} du \\ &+ \mathbb{E}^{\mathbb{Q}^d} \left[\int_0^t \mathbf{1}_{S_u \geq K} D_u^d \alpha(S_u, u) S_u \sqrt{V_u} dW_u \right]. \end{aligned} \quad (3.10)$$

Furthermore, on a fixed time interval $[0, T^*]$, D_t^d is uniformly bounded by $e^{T^* \max_{u \in [0, T^*]} |h^d(u)|}$. Then, from Lemma 3 we know that $\int_0^t \mathbf{1}_{S_u \geq K} D_u^d \alpha(S_u, u) S_u \sqrt{V_u} dW_u$ is a true martingale of zero expectation.

We write (3.10) at time T , differentiate with respect to T and, upon noticing that $\mathbf{1}_{S_T \geq K} S_T = (S_T - K)^+ + \mathbf{1}_{S_T < K} K$, we get (3.6). \square

Proof of Theorem 1. First, we want to ensure that (3.1) is a necessary condition for

$$C(K, T) = C_{LV}(K, T). \quad (3.11)$$

Hence, by subtracting the Dupire PDE (2.4) from (3.6), we obtain

$$\begin{aligned} \frac{1}{2} K^2 \left(\alpha(K, T)^2 \frac{\mathbb{E}^{\mathbb{Q}^d} [D_T^d V_T | S_T = K]}{\mathbb{E}^{\mathbb{Q}^d} [D_T^d | S_T = K]} - \sigma_{LV}(K, T)^2 \right) \frac{\partial^2 C_{LV}}{\partial K^2} &= \mathbb{E}^{\mathbb{Q}^d} [D_T^d r_T^f (S_T - K)^+] \\ &- \bar{r}^f(T) C_{LV} - \mathbb{E}^{\mathbb{Q}^d} \left[D_T^d \mathbf{1}_{S_T \geq K} K \left(r_T^d - r_T^f \right) \right] - K (\bar{r}^d(T) - \bar{r}^f(T)) \frac{\partial C_{LV}}{\partial K}, \end{aligned}$$

so

$$\alpha^2(K, T) = \frac{\mathbb{E}^{\mathbb{Q}^d} [D_T^d | S_T = K]}{\mathbb{E}^{\mathbb{Q}^d} [D_T^d V_T | S_T = K]} \left(\sigma_{LV}(K, T)^2 + \bar{q}(K, T) \right),$$

where

$$\bar{q}(K, T) = \frac{\mathbb{E}^{\mathbb{Q}^d} [\bar{Q}_T]}{\frac{1}{2} K^2 \frac{\partial^2 C_{LV}}{\partial K^2}},$$

$$\bar{Q}_T = D_T^d r_T^f (S_T - K)^+ - \bar{r}^f(T) C_{LV} - K \left(D_T^d \mathbf{1}_{S_T \geq K} (r_T^d - r_T^f) + (\bar{r}^d(T) - \bar{r}^f(T)) \frac{\partial C_{LV}}{\partial K} \right).$$

It remains to show that $\bar{q} = q$. First, $D_T^d (S_T - K)^+$ is weakly differentiable with respect to K with $\frac{\partial D_T^d (S_T - K)^+}{\partial K} = D_T^d \mathbf{1}_{S_T \geq K}$, which is bounded by the integrable process D_T^d . We can interchange differentiation and expectation to get $\mathbb{E}^{\mathbb{Q}^d} [D_T^d \mathbf{1}_{S_T \geq K}] = -\frac{\partial C}{\partial K}$. Since the models agree, $\mathbb{E}^{\mathbb{Q}^d} [D_T^d \mathbf{1}_{S_T \geq K}] = -\frac{\partial C_{LV}}{\partial K}$ and $\mathbb{E}^{\mathbb{Q}^d} [D_T^d (S_T - K)^+] = C_{LV}$, and (3.1) holds. Moreover, it is easy to see that (3.1) is also a sufficient condition for (3.11). \square

Remark. In this setting, we rely on the fact that a pure LV model, as defined in (2.3), can reproduce any arbitrage-free market call prices. In [19], no-arbitrage constraints on call prices are given. Among others, one needs price convexity (to ensure existence of a non-negative risk-neutral density by the Breeden and Litzenberger formula), and calendar spreads to be positive. Namely for a market call price C and all $0 < T_1 < T_2$,

$$\begin{aligned} \frac{\partial^2 C}{\partial K^2} &\geq 0, \\ C \left(K e^{\int_{T_1}^{T_2} (\bar{r}^d(t) - \bar{r}^f(t)) dt}, T_2 \right) &\geq e^{-\int_{T_1}^{T_2} \bar{r}^f(t) dt} C(K, T_1). \end{aligned} \quad (3.12)$$

In [9], it is shown that an equivalent formulation of (3.12) is

$$\frac{\partial C(K, T)}{\partial T} + \bar{r}^f(T) C(K, T) + K (\bar{r}^d(T) - \bar{r}^f(T)) \frac{\partial C(K, T)}{\partial K} \geq 0.$$

Additionally, the Dupire formula

$$\sigma_{LV}^2(K, T)^2 = \frac{\frac{\partial C(K, T)}{\partial T} + \bar{r}^f(T) C(K, T) + K (\bar{r}^d(T) - \bar{r}^f(T)) \frac{\partial C(K, T)}{\partial K}}{\frac{1}{2} K^2 \frac{\partial^2 C(K, T)}{\partial K^2}} \quad (3.13)$$

provides a necessary and sufficient condition for the pure LV model (2.3) to match market call prices. Existence of a local volatility function is then guaranteed by the (static) no-arbitrage conditions that preclude the numerator or denominator to become negative in (3.13) (see also [8] for details).

We also note that for a fixed set of stochastic rates parameters, q can become negative and potentially larger than σ_{LV}^2 in (3.1). In that case, there exists no leverage function α such that model (2.1) fits the market call prices. However, in practice we did not encounter any problem by calibrating the stochastic rates parameters directly to the market. Additionally, since model (2.1) is a generalisation of (2.3), there always exists a set of stochastic rates parameters such that α exists and satisfies (3.1).

4 Fast calibration with a new control variate particle method

In this section we describe the full calibration routine step-by-step. First, we give the market data. Then we present the various control variates that we use in order to drastically reduce the computational cost of the particle method for α in the 4-factor model (5.1), which build on intermediary calibration steps. Second, we detail each of these steps, including calibration of the LV, LSV, and interest rate models.

4.1 Market data set

We use a set of market data from Bloomberg from the 18/03/2016 for the currency pair EURUSD. The spot value is denominated as the price of one EUR in USD units,

$$\text{Spot} = 1.1271 \text{ (\$ per €) ,}$$

and the implied volatility is retrieved for the Deltas

$$10\text{D-Put}, 25\text{D-Put}, 50\text{D}, 25\text{D-Call}, 10\text{D-Call}$$

and for the maturity pillars

$$3\text{W}, 1\text{M}, 2\text{M}, 3\text{M}, 6\text{M}, 1\text{Y}, 1\text{Y}6\text{M}, 2\text{Y}, 3\text{Y}, 5\text{Y}.$$

As an aside, we skip the 7Y and 10Y pillars as the quotes we retrieved were not liquid enough. We plot the market volatility surface on a strike scale in Figure 4.1 (from 6 months onward for a better illustration of the shape). The volatility is interpolated linearly in variance on a log-moneyness scale and extrapolated linearly.

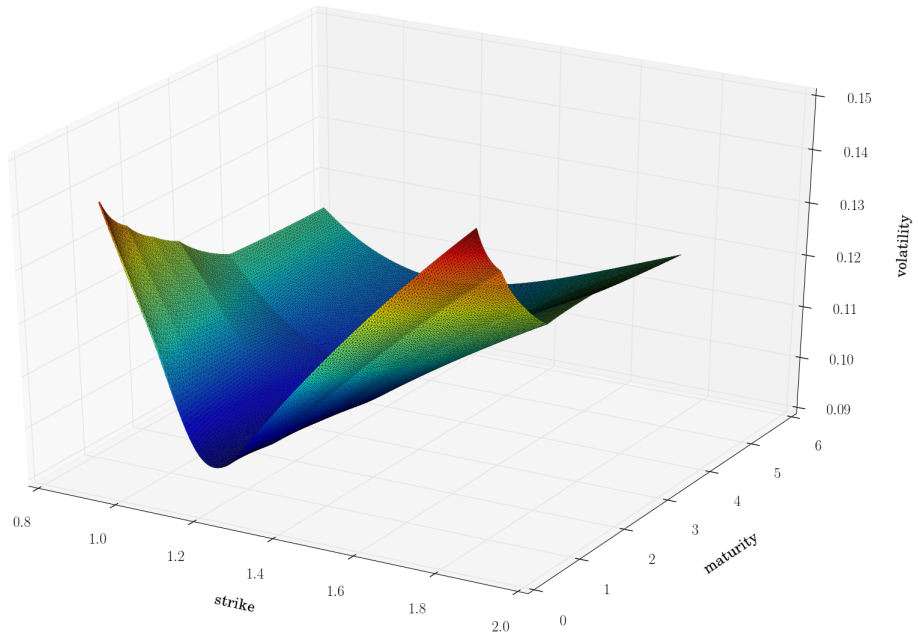


Figure 4.1: Market volatility surface for the EUR-USD (18/03/2016)

The domestic risk-neutral measure \mathbb{Q}^d is the one associated with the USD money market account. The foreign risk-neutral measure \mathbb{Q}^f is the one associated to the EUR money market account. The net present value of any financial contract will be expressed in units of domestic numéraire unless stated otherwise.

4.2 Variance reduction for standard expectations

If we compare formula (3.4) and the formula in Theorem 1, we notice that when the rates are stochastic, an additional term appears in the leverage function, namely

$$q(K, T) = \frac{\mathbb{E}^{\mathbb{Q}^d}[Q_T]}{\frac{1}{2}K^2 \frac{\partial^2 C_{LV}}{\partial K^2}},$$

where

$$Q_T = \underbrace{D_T^d \left(r_T^f - \bar{r}^f(T) \right) (S_T - K)^+}_{=: X_T} - K \underbrace{D_T^d \mathbf{1}_{S_T \geq K} \left[(r_T^d - \bar{r}^d(T)) - (r_T^f - \bar{r}^f(T)) \right]}_{=: X'_T}. \quad (4.1)$$

In this section we assume that a pure local volatility model (2.3) has been previously calibrated to market call option prices such that for any quoted strike K and maturity T ,

$$C_{LV}(K, T) = C_{Market}(K, T).$$

First, we note that the function $\frac{\partial^2 C_{LV}}{\partial K^2}$ can become very small, which can lead to numerical instabilities. To the best of our knowledge, in the literature ([25, 38, 10]) the Dupire formula is typically used to compute this quantity, basically, by differentiating the market prices twice. However, this approach leads to instabilities as a small error in the computation of this second derivative can lead to a big difference in the estimated value of q . Therefore, we propose to calibrate a pure local volatility model with the forward PDE as explained in Subsection 4.6 and use $\frac{\partial^2 C_{LV}}{\partial K^2}$ obtained from the PDE solution with a smoothing scheme. Then this quantity will be smooth and stable.

Second, for the random variable Q and for an associated N -sample $(Q_i)_{i \in [1, N]}$, we use the sample mean $\frac{1}{N} \sum_{i=1}^N Q_i$ as the Monte Carlo estimator of $\mathbb{E}^{\mathbb{Q}^d}[Q]$ with control variates

$$Y_T = D_T^d (S_T - K)^+, \quad Z_T = r_T^f - \bar{r}^f(T)$$

for X_T defined in (4.1), and

$$Y'_T = D_T^d \mathbf{1}_{S_T \geq K}, \quad Z'_T = (r_T^d - \bar{r}^d(T)) - (r_T^f - \bar{r}^f(T))$$

for X'_T . We know that the model (2.1) is perfectly calibrated to call option prices if and only if

$$\begin{aligned} \mu_1 &:= \mathbb{E}^{\mathbb{Q}^d}[Y_T] = C_{LV}(K, T), \\ \mu_2 &:= \mathbb{E}^{\mathbb{Q}^d}[Y'_T] = -\frac{\partial C_{LV}}{\partial K}(K, T). \end{aligned}$$

Let us also define

$$\gamma_1 := \mathbb{E}^{\mathbb{Q}^d}[X_T], \quad \gamma_2 := \mathbb{E}^{\mathbb{Q}^d}[X'_T]$$

and

$$\begin{aligned} \zeta_1 &:= \mathbb{E}^{\mathbb{Q}^d}[Z_T] = g_0^f e^{-\kappa_f T} + \theta_f (1 - e^{-\kappa_f T}) + h^f(T) - \bar{r}^f(T), \\ \zeta_2 &:= \mathbb{E}^{\mathbb{Q}^d}[Z'_T] = g_0^d e^{-\kappa_d T} + \theta_d (1 - e^{-\kappa_d T}) + h^d(T) - \bar{r}^d(T) - \zeta_1. \end{aligned}$$

We denote the corresponding Monte Carlo estimators by $\hat{\mu}_{1,N}$, $\hat{\mu}_{2,N}$, $\hat{\gamma}_{1,N}$, $\hat{\gamma}_{2,N}$, $\hat{\zeta}_{1,N}$ and $\hat{\zeta}_{2,N}$, respectively, using the same Brownian paths for W , W^V , W^{g^d} , W^{g^f} in all estimators. We can define a new Monte Carlo estimator γ_N^* for $\gamma_1 - K\gamma_2$,

$$\gamma_N^* = (\gamma_{1,N}^* - K\gamma_{2,N}^*), \quad (4.2)$$

with

$$\begin{aligned} \gamma_{1,N}^* &= \hat{\gamma}_{1,N} + \lambda_1 (\hat{\mu}_{1,N} - C_{LV}(K, T)) + \eta_1 (\hat{\zeta}_{1,N} - \zeta_1), \\ \gamma_{2,N}^* &= \hat{\gamma}_{2,N} + \lambda_2 \left(\hat{\mu}_{2,N} + \frac{\partial C_{LV}(K, T)}{\partial K} \right) + \eta_2 (\hat{\zeta}_{2,N} - \zeta_2). \end{aligned}$$

We define a random variable

$$G_T = X_T + \lambda_1 (Y_T - C_{LV}(K, T)) + \eta_1 (Z_T - \zeta_1)$$

such that $\gamma_{1,N}^*$ is an estimator of $\mathbb{E}[G_T] = \mathbb{E}[X_T]$, and we similarly define a random variable G' involving Y' and Z' , with an estimator $\gamma_{2,N}^*$. Each expectation now has two control variates (see Section 4.1 in [22]). The variance of G_T is

$$\begin{aligned} \mathbb{V}[X_T + \lambda_1(Y_T - C_{LV}(K, T)) + \eta_1(Z_T - \zeta_1)] &= \mathbb{V}[X_T] + \lambda_1^2 \mathbb{V}[Y_T] + \eta_1^2 \mathbb{V}[Z_T] \\ &\quad + 2\lambda_1 \text{Cov}(X_T, Y_T) + 2\eta_1 \text{Cov}(X_T, Z_T) + 2\lambda_1 \eta_1 \text{Cov}(Y_T, Z_T). \end{aligned}$$

A minimum is obtained (see [22]) when $\lambda_1 = \bar{\lambda}(X_T, Y_T, Z_T)$ and $\eta_1 = \bar{\lambda}(X_T, Z_T, Y_T)$, where

$$\bar{\lambda}(X, Y, Z) = \frac{\text{Cov}(X, Y) \text{Cov}(Y, Z) - \text{Cov}(X, Y) \mathbb{V}[Z]}{\mathbb{V}[Z] \mathbb{V}[Y] - \text{Cov}(Y, Z)^2},$$

which can be estimated from the Monte Carlo samples. Similar expressions are obtained for X' , where the corresponding optimisers are $\lambda_2 = \bar{\lambda}(X'_T, Y'_T, Z'_T)$, $\eta_2 = \bar{\lambda}(X'_T, Z'_T, Y'_T)$. Then

$$q(K, T) = \lim_{N \rightarrow \infty} \frac{\gamma_N^*}{\frac{1}{2} K^2 \frac{\partial^2 C_{LV}}{\partial K^2}}.$$

This approach is particularly useful for out-of-the-money options and digital options as the Monte Carlo estimator will exhibit higher variance in these settings. The method will help reduce the variance while working with a sufficiently low number of particles (i.e., paths).

One should be careful in the eventuality where there is no particle S_T above K . In that case, the estimated variances $\mathbb{V}[D_T^d \mathbf{1}_{S_T \geq K}]$ and $\mathbb{V}[D_T^d (S_T - K)^+]$ are zero, and we pick

$$\begin{aligned} \lambda_1 &= -\zeta_1, \quad \eta_1 = -C_{LV}(K, T), \\ \lambda_2 &= -\zeta_2, \quad \eta_2 = \frac{\partial C_{LV}(K, T)}{\partial K}, \end{aligned}$$

such that both control variates are of the same order of magnitude. Similarly, if all particles S_T are above K , the estimated variance $\mathbb{V}[D_T^d \mathbf{1}_{S_T \geq K}]$ is zero. We then choose

$$\lambda_2 = -\zeta_2, \quad \eta_2 = \frac{\partial C_{LV}(K, T)}{\partial K}.$$

4.3 Variance reduction for the Markovian projection

Our goal is to reduce the variance of an estimator p_N for

$$p(K, T) = \frac{\mathbb{E}^{\mathbb{Q}^d}[D_T^d V_T | S_T = K]}{\mathbb{E}^{\mathbb{Q}^d}[D_T^d | S_T = K]}.$$

Let us assume that we have simulated N Monte Carlo paths for the system (D_t^d, S_t, V_t) up to time $T > 0$. We denote the state of this system at time T for the i -th path by (D_T^i, S_T^i, V_T^i) . Following the particle method used for LSV model calibration in [25], an estimator for p is defined as

$$\begin{aligned} \hat{p}_N(K, T) &= \frac{\sum_{i=1}^N D_T^i V_T^i \delta_N(S_T^i - K)}{\sum_{i=1}^N \delta_N(S_T^i - K)} \frac{\sum_{i=1}^N \delta_N(S_T^i - K)}{\sum_{i=1}^N D_T^i \delta_N(S_T^i - K)} \\ &= \frac{\sum_{i=1}^N D_T^i V_T^i \delta_N(S_T^i - K)}{\sum_{i=1}^N D_T^i \delta_N(S_T^i - K)}, \end{aligned}$$

where δ_N is a kernel function with a bandwidth that “tightens” with the number of particles. We assume that the 2-factor LSV model (2.2) is perfectly calibrated to market call prices. We know from [38, 18, 25] that this is the case if and only if

$$\alpha^{2D}(K, T) = \frac{\sigma_{LV}(K, T)}{\sqrt{\mathbb{E}^{\mathbb{Q}^d}[V_T^{2D} | S_T^{2D} = K]}}.$$

We can use

$$p^{2D} = \mathbb{E}^{\mathbb{Q}^d} [V_T^{2D} | S_T^{2D} = K]$$

and its Monte Carlo estimator,

$$\hat{p}_N^{2D}(K, T) = \sum_{i=1}^N \frac{V_T^{2D,i} \delta_N(S_T^{2D,i} - K)}{\sum_{i=1}^N \delta_N(S_T^{2D,i} - K)}, \quad (4.3)$$

as a control variate for \hat{p}_N . We define a new estimator p_N^* such that

$$p_N^* = \hat{p}_N(K, T) + \lambda (\hat{p}_N^{2D}(K, T) - p^{2D}(K, T)). \quad (4.4)$$

The latter has an asymptotically diminishing bias if we assume the particle method to converge in distribution (and neglect the time stepping bias).

In order to get a good estimate for λ , we can rewrite the above estimator as follows

$$p_N^* = \frac{1}{N} \sum_{i=1}^N \frac{D_T^i V_T^i \delta_N(S_T^i - K)}{\frac{1}{N} \sum_{i=1}^N D_T^i \delta_N(S_T^i - K)} + \lambda \left(\frac{1}{N} \sum_{i=1}^N \frac{V_T^{2D,i} \delta_N(S_T^{2D,i} - K)}{\frac{1}{N} \sum_{i=1}^N \delta_N(S_T^{2D,i} - K)} - p^{2D}(K, T) \right),$$

which can be also written as

$$p_N^* = \frac{1}{N} \sum_{i=1}^N m_i + \lambda \left(\frac{1}{N} \sum_{i=1}^N m_i^{2D} - p^{2D}(K, T) \right)$$

with

$$m_i = \frac{D_T^i V_T^i \delta_N(S_T^i - K)}{\frac{1}{N} \sum_{i=1}^N D_T^i \delta_N(S_T^i - K)}, \quad m_i^{2D} = \frac{V_T^{2D,i} \delta_N(S_T^{2D,i} - K)}{\frac{1}{N} \sum_{i=1}^N \delta_N(S_T^{2D,i} - K)}$$

which is in the standard Monte Carlo control variate form. We can think of m_i and m_i^{2D} as samples of two random variables m and m^{2D} respectively, and for the best variance reduction (see Section 4.1 in [22]), we take

$$\lambda = -\frac{\text{Cov}(m, m^{2D})}{\text{Var}(m^{2D})},$$

which we can estimate by

$$\hat{\lambda}_N = -\frac{\sum_{i=1}^N (m_i - \bar{m})(m_i^{2D} - \bar{m}^{2D})}{\sum_{i=1}^N (m_i^{2D} - \bar{m}^{2D})^2}, \quad (4.5)$$

where \bar{m} and \bar{m}^{2D} are the sample means. We recall that the expected variance reduction factor is

$$\frac{1}{1 - \text{Corr}(m, m^{2D})^2}. \quad (4.6)$$

Hence, if the stochastic rates are not highly volatile, i.e. if ξ_f and ξ_d are small enough, we expect a very good variance reduction as the correlation between the particles generated by the 4-factor hybrid LSV model (2.1) and by the 2-factor LSV model (2.2) will be high. Our numerical tests performed on a model calibrated to recent EUR and USD market data exhibit a correlation (measured for random variables m and m^{2D}) between model (2.1) and (2.2) of 95% up to 1.5 years and 50% around 5 years. Additionally, our stress test in Subsection 5.3 suggests that even under high volatility regimes for the rate processes, i.e. when ξ_f and ξ_d are very big, the variance reduction brought by this control variate is significant.

4.4 Shifted CIR model

The domestic and foreign short interest rates are modeled by the shifted CIR (CIR++) process [6]. On the one hand, this model preserves the analytical tractability of the CIR model for bonds, caps and other basic interest rate products. On the other hand, it is flexible enough to fit the initial term structure of interest rates exactly. For $i \in \{d, f\}$, the short rate dynamics under their respective spot measures, i.e., \mathbb{Q}^d – domestic and \mathbb{Q}^f – foreign, are given by

$$\begin{cases} r_t^i = g_t^i + h^i(t), \\ dg_t^i = \kappa_i(\theta_i - g_t^i)dt + \xi_i \sqrt{g_t^i} dB_t^{g^i}, \quad g_0^i > 0, \end{cases} \quad (4.7)$$

where B^{g^d} and B^{g^f} are Brownian motions under \mathbb{Q}^d and \mathbb{Q}^f , respectively. The mean-reversion parameters κ_i , the long-term mean parameters θ_i and the volatility parameters ξ_i are the same as in (2.1). The calibration of the short rate model (4.7) follows the same approach for both the domestic and the foreign interest rate. For simplicity, we drop the subscripts and superscripts “ d ” and “ f ” in the remainder of the subsection and define the vector of parameters $\beta_1 = (g_0, \kappa, \theta, \xi)$. According to Brigo and Mercurio [5], an exact fit to the initial term structure of interest rates is equivalent to $h(t) = \varphi^{\text{CIR}}(t; \beta_1)$ for all $t \in [0, T]$, where

$$\begin{aligned} \varphi^{\text{CIR}}(t; \beta_1) &= \bar{r}(0, t) - \bar{r}^{\text{CIR}}(0, t; \beta_1), \\ \bar{r}^{\text{CIR}}(0, t; \beta_1) &= \frac{2\kappa\theta(\exp\{t\nu\} - 1)}{2\nu + (\kappa + \nu)(\exp\{t\nu\} - 1)} + g_0 \frac{4\nu^2 \exp\{t\nu\}}{[2\nu + (\kappa + \nu)(\exp\{t\nu\} - 1)]^2}, \end{aligned} \quad (4.8)$$

$\nu = \sqrt{\kappa^2 + 2\xi^2}$ and $\bar{r}(0, t)$ is the market instantaneous forward rate at time 0 for a maturity t , i.e.,

$$\bar{r}(0, t) = \bar{r}(t) = -\frac{\partial \ln P(0, t)}{\partial t}, \quad (4.9)$$

where $P(0, t)$ is the market zero coupon bond price at time 0 for a maturity t . The value of the zero coupon bond is given by

$$P(0, t) = \frac{1}{1 + \Delta(0, t)R(0, t)}, \quad (4.10)$$

where $\Delta(0, t)$ is the year fraction from 0 to time t and $R(0, t)$ is the current (simply-compounded) deposit rate with maturity date t which is quoted in the market. As an aside, note that the standard day count convention for USD and EUR is Actual 360.

The detailed calibration procedure for both domestic and foreign rate processes can be found in Appendix A.3. The calibration results are displayed in Table 1.

Table 1: The calibrated CIR parameters

CCY	g_0	κ	θ	ξ
USD	0.0001	0.0837	0.5469	0.0274
EUR	0.0001	0.0110	1.1656	0.0370

4.5 Four-factor hybrid stochastic volatility model calibration

Consider a “purely stochastic” version of the model (2.1) – the Heston-2CIR++ model with leverage function $\equiv 1$ – and additionally suppose that the domestic and the foreign short interest rate dynamics are independent of the dynamics of the spot FX rate. The model is governed by the

following system of SDEs under the domestic risk-neutral measure \mathbb{Q}^d :

$$\begin{cases} \frac{dS_t^{SV}}{S_t^{SV}} = (r_t^d - r_t^f) dt + \sqrt{V_t} dW_t, & S_0^{SV} = S_0, \\ r_t^d = g_t^d + h^d(t) \\ r_t^f = g_t^f + h^f(t) \\ dg_t^d = \kappa_d (\theta_d - g_t^d) dt + \xi_d \sqrt{g_t^d} dW_t^{g^d} \\ dg_t^f = \kappa_f (\theta_f - g_t^f) dt + \xi_f \sqrt{g_t^f} dW_t^{g^f} \\ dV_t = \kappa (\theta - V_t) dt + \xi \sqrt{V_t} dW_t^V, \end{cases} \quad (4.11)$$

where W^s and W^V are correlated Brownian motions with correlation coefficient ρ . Note that the quanto correction term in the drift of the foreign short rate vanishes due to the postulated independence assumption between the spot FX rate and foreign short rate dynamics.

Define the vector of parameters $\beta_2 = (v_0, \kappa, \theta, \xi, \rho)$. The next step in our calibration is to find the values of these 5 model parameters for which European call option prices best match the market call prices retrieved from volatility quotes for different strikes and maturities. For a EURUSD transaction, the market standard is to choose USD as the domestic currency and EUR as the foreign currency. The forward FX rate for a payment date T is defined as

$$F_T = \frac{P^f(0, T)}{P^d(0, T)} S_0, \quad (4.12)$$

where $P^d(0, T)$ and $P^f(0, T)$ are the domestic and foreign discount factors at time 0 for a maturity T , respectively.

Under the postulated simple correlation structure of the Brownian drivers and when the short rates are driven by the CIR process, i.e., when $h^{d,f} = 0$, Ahlip and Rutkowski [2] derive an efficient closed-form formula for the European call option price. Hence, we denote by $C_A(K, T)$ the fair value under the Heston-2CIR model of a European call option with strike K and maturity T computed with the aforementioned formula, and by $C_H(K, T)$ the fair value of the same option but under the Heston-2CIR++ model. For $i \in \{d, f\}$, we define for brevity

$$H_i = \exp \left\{ \int_0^T h^i(t) dt \right\},$$

where the shift functions $h^{d,f}$ were calibrated in the previous subsection. Then we can extend the pricing formula of Ahlip and Rutkowski [2] as follows.

$$\begin{aligned} C_H(K, T) &= \mathbb{E}^{\mathbb{Q}^d} \left[\exp \left\{ - \int_0^T r_t^d dt \right\} (S_T^{SV} - K)^+ \right] \\ &= H_f^{-1} \mathbb{E}^{\mathbb{Q}^d} \left[\exp \left\{ - \int_0^T g_t^d dt \right\} (H_f H_d^{-1} S_T^{SV} - H_f H_d^{-1} K)^+ \right]. \end{aligned}$$

Therefore, $C_H(K, T) = H_f^{-1} C_A(\tilde{K}, T)$, where $\tilde{K} = H_f H_d^{-1} K$. We now calibrate the Heston-2CIR++ model by minimising the sum of the squared differences between model and market call prices:

$$\min_{\beta_2 \in \mathbb{R}_+^4 \times [-1, 1]} \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}} [C_H(K_j, T_i; \beta_2) - C_{BS}(K_j, T_i, \sigma_{i,j})]^2, \quad (4.13)$$

where $\sigma_{i,j}$ is the quoted volatility corresponding to a strike K_j and a maturity T_i , for $j = 1, \dots, m$ and $i = 1, \dots, n$. There are many ways to choose the objective function (error measure) in (4.13). For instance, we may consider either call prices or Black-Scholes implied volatilities and minimise

the sum of either absolute or relative (squared) differences between model and market values, using either uniform or non-uniform weights. We choose this particular error measure, which assigns more weight to more expensive options (in-the-money, long-term) and less weight to cheaper options (out-of-the-money, short-term), for two reasons. First, the Heston model, and hence the Heston-2CIR++ model by extension, cannot reproduce the smiles or skews typically observed for short maturities that well and a more careful calibration to these smiles would result in a larger overall model error due to the inherent poor fit of the model to the short-term. Second, market data becomes scarce as the maturity increases, and hence we already assigned more weight to the short- and mid-term sections of the volatility surface; for instance, we have more maturities up to 1 year than between 1 and 5 years. The calibration results are displayed in Table 2.

Table 2: The calibrated Heston parameters

v_0	κ	θ	ξ	ρ
0.0094	1.4124	0.0137	0.2988	-0.1194

As before, we employ a nonlinear least-squares solver (the trust-region-reflective algorithm, see [11]) for the calibration and a global optimisation method (a genetic algorithm) for verification purposes. Due to the non-linearity and non-convexity of the problem, the calibrated model parameters may end up in a local rather than a global minimum of the objective function. Hence, a good initial parameter guess may significantly improve the quality of the calibration. Practitioners usually use variance swap prices to calibrate v_0 , κ and θ . In our case, we found the squared ATM 3-week and 5-year volatilities to provide good initial guesses for v_0 and θ , respectively.

4.6 One-factor local volatility model calibration

The calibration routine for a pure local volatility model is run with a standard algorithm forward in maturity. We recall that model (2.3) is written as

$$\frac{dS_t^{LV}}{S_t^{LV}} = (\bar{r}^d(t) - \bar{r}^f(t)) dt + \sigma_{LV}(S_t^{LV}, t) dW_t,$$

and we want to find the function σ_{LV} for which the call prices under the local volatility model match the quoted market prices exactly. This is crucial as both σ_{LV} and $\frac{\partial^2 C_{LV}}{\partial K^2}$ appear in the leverage function formula (3.1). The forward Dupire PDE (2.4),

$$\begin{cases} \frac{\partial C_{LV}}{\partial T} + (\bar{r}^d(T) - \bar{r}^f(T)) K \frac{\partial C_{LV}}{\partial K} + \bar{r}^f(T) C_{LV} - \frac{1}{2} K^2 \frac{\partial^2 C_{LV}}{\partial K^2} \sigma_{LV}^2(K, T) = 0, \\ C_{LV}(K, 0) = (S_0 - K)^+, \quad C_{LV}(0, T) = S_0, \quad C_{LV}(S_{max}, T) = 0, \end{cases} \quad (4.14)$$

provides an efficient way to calibrate and, eventually, regularise the problem. Denote by Φ the map from the local volatility to the model implied volatility Σ_{Model} , i.e.,

$$\Phi(\{\sigma_{LV}(s, t) \mid (s, t) \in \mathbb{R}_+^2\}) = \{\Sigma(K, T) \mid (K, T) \in \mathbb{R}_+^2\}.$$

Furthermore, the PDE solution for a guess σ_S of the local volatility gives call prices for the whole set of strikes and maturities. Inverting the Black formula allows to retrieve the model implied volatilities Σ . Hence, as proposed in [45], we can use the forward Dupire PDE (4.14) combined with an efficient implied volatility inverter [29] as the mapping function Φ . A very useful property of this PDE is that it can be solved forward in maturity. Let a set of maturities quoted on the market be $(T_1, \dots, T_{N_{Mat}})$ and a set of M_i strikes for a given maturity T_i be $(K_{T_i,1}, \dots, K_{T_i,M_i})$. It is possible to solve the PDE on $[0, T_1]$, then on $[T_1, T_2]$ and so forth. The full calibration algorithm is presented for completeness in Appendix A.2.

4.7 Two-factor Heston-type LSV model calibration by PDE

The next step in our algorithm is to calibrate the 2-factor sub-model of (2.1) defined in (2.2), which we recall below. In this subsection only, we write S, V and α in lieu of S^{2D}, V^{2D} and α^{2D} , respectively, to ease notation. Hence,

$$\begin{cases} \frac{dS_t}{S_t} = (\bar{r}^d(t) - \bar{r}^f(t)) dt + \alpha(S_t, t) \sqrt{V_t} dW_t, \\ dV_t = \kappa(\theta - V_t) dt + \xi \sqrt{V_t} dW_t^V. \end{cases}$$

We use the Heston parameters calibrated in Subsection 4.5, i.e.,

$$v_0 = 0.0094, \quad \theta = 0.0137, \quad \kappa = 1.4124, \quad \rho = -0.1194, \quad \xi = 0.2988,$$

where the Feller ratio is

$$\frac{2\kappa\theta}{\xi^2} \approx 0.4335 < 1,$$

which clearly breaks the Feller condition. The calibration of these models by PDEs is extensively discussed in [38, 10, 52] and the approach proves to be very effective when the variance process has almost surely strictly positive paths. However, for Heston type models, the CIR variance process can eventually reach zero if the Feller condition is violated. And as a matter of fact, in FX markets, this condition is very often violated. One can refer to Table 6.5 in [10] for examples on a large range of currency pairs and maturities. As a consequence, mass of the variance process will accumulate towards $V = 0$ and the density will explode at this boundary. In [44] the authors propose to reduce the problem by considering $\log(V_t/v_0)$, whereas [10] suggests to refine the mesh near $V = 0$. While these methods can provide acceptable numerical results, they do not entirely solve the problem since $\log(V_t/v_0)$ can reach $-\infty$ with non-zero probability. Hence, in this work, we propose to use a different boundary condition as well as a change of variable to control the solution in the neighbourhood of $V = 0$ such that the density is finite on the boundary.

Transformation and weak formulation

The following is a small variation of the main result in [30], and the proof is therefore omitted. Note that a new non-Dirichlet boundary condition appears at $z = 0$.

Theorem 5. *Define the region $\Omega = \{(x, z) \in \mathbb{R}_+^2\}$ and assume that the density ϕ (under \mathbb{Q}^d) of the Markovian process (S_t, V_t) started at (S_0, v_0) at time 0 exists and is $C^{1,2,2}(\mathbb{R}^+ \times \Omega)$. Then ϕ is the solution to the Kolmogorov forward equation*

$$\begin{cases} \frac{\partial \phi}{\partial t} + (\bar{r}^d(t) - \bar{r}^f(t)) \frac{\partial x \phi}{\partial x} + \frac{\partial \kappa(\theta - z) \phi}{\partial z} \\ \quad - \frac{1}{2} \left(\frac{\partial^2 x^2 \alpha^2(x, t) z \phi}{\partial x^2} + \frac{\partial^2 \xi^2 z \phi}{\partial z^2} + 2 \frac{\partial^2 \rho \xi x \alpha(x, t) z \phi}{\partial x \partial z} \right) = 0, & (x, z) \in \Omega, t > 0, \\ \left(\frac{\xi^2}{2} \frac{\partial z \phi}{\partial z} - \kappa(\theta - z) \phi + \rho \xi z \frac{\partial x \alpha(x, t) \phi}{\partial x} \right) \Big|_{z=0} = 0, & z = 0, x \geq 0, t > 0, \\ \lim_{z \rightarrow \infty} \phi(x, z, t) = \lim_{x \rightarrow \infty} \phi(x, z, t) = \phi(0, z, t) = 0, & (x, z) \in \Omega, t > 0, \\ \lim_{t \rightarrow 0} \phi(x, y, z, t) = \delta(x - S_0, z - v_0), & (x, z) \in \Omega. \end{cases} \quad (4.15)$$

Proof. Similar to the proof of Lemma 4.1 and Theorem 4.1 in [30]. \square

The density function for a CIR process is written as follows (see [13]),

$$\phi_v(z, t) = c e^{-u - cz} \left(\frac{cz}{u} \right)^{\beta/2} I_\beta(2\sqrt{cu}z),$$

with $c = \frac{2\kappa}{(1 - e^{-\kappa t})\xi^2}$, $u = cv_0 e^{-\kappa t}$ and $\beta = \frac{2\kappa\theta}{\xi^2} - 1$, where I_β is the modified Bessel function of the first kind of order β . We can write the asymptotic expression for small z using the asymptotic formula for the modified Bessel function found in [1]. We get, for z around 0^+ ,

$$\phi_v(z, t) \sim \frac{c^{(\beta+1)} e^{-u - cz}}{\Gamma(\beta + 1)} z^\beta, \quad (4.16)$$

such that $\phi_v(z, t)$ diverges for $z = 0^+$ when $2\kappa\theta < \xi^2$. This is in line with the well-known stationary distribution density ϕ_v^∞ (see [13]) defined as

$$\phi_v^\infty(z) = \lim_{t \rightarrow \infty} \phi_v(z, t) = \frac{\omega^{(\beta+1)} z^\beta e^{-\omega z}}{\Gamma(\beta+1)}, \quad (4.17)$$

with $\omega = \frac{2\kappa}{\xi^2}$. This justifies a change of variable for the density ϕ of the form $p = \phi z^{-\beta}$, and to solve a new PDE for p that we define hereafter. We can insert $z^\beta p$ and its associated derivatives in (4.15). By straightforward calculus we get the following.

Corollary 6. *The dampened “density” $p = \phi z^{-\beta}$ satisfies the initial boundary value problem*

$$\begin{cases} \frac{\partial p}{\partial t} + (\bar{r}^d(t) - \bar{r}^f(t)) \frac{\partial x p}{\partial x} - \beta \kappa p \\ \quad + \frac{\partial \kappa(\theta - z)p}{\partial z} - (\beta + 1) \left(\frac{\partial \xi^2 p}{\partial z} + \frac{\partial \rho \xi x \alpha(x, t) p}{\partial x} \right) \\ \quad - \frac{1}{2} z \left[\frac{\partial^2 x^2 \alpha^2(x, t) p}{\partial x^2} + \frac{\partial^2 \xi^2 p}{\partial z^2} + 2 \frac{\partial^2 \rho \xi x \alpha(x, t) p}{\partial x \partial z} \right] = 0, & (x, z) \in \Omega, t > 0, \\ \left[\frac{\xi^2 z}{2} \frac{\partial p}{\partial z} \right]_{z=0} + \kappa z p \Big|_{z=0} + \rho \xi z \frac{\partial x \alpha(x, t) p}{\partial x} \Big|_{z=0} = 0, & z = 0, t > 0, \\ \lim_{z \rightarrow \infty} p(x, z, t) = \lim_{x \rightarrow \infty} p(x, z, t) = p(0, z, t) = 0, & (x, z) \in \Omega, z \neq 0, t > 0, \\ \lim_{t \rightarrow 0} p(x, z, t) = z^{-\beta} \delta(x - S_0, z - v_0), & (x, z) \in \Omega, \end{cases} \quad (4.18)$$

with $\beta = \frac{2\kappa\theta}{\xi^2} - 1$.

While this PDE is easier to handle numerically, one wants to work with the original density function ϕ for most of the applications. There are two main calculations one would like to achieve:

- compute the price of a derivative with final payoff value $f(S_T)$ for a given function f as

$$\mathbb{E}[D_T f(S_T)] = D_T \int_0^\infty f(x) \int_0^\infty \phi(x, z, T) dz dx;$$

- compute the Markovian projection

$$\mathbb{E}[V_T | S_T = K] = \frac{\int_0^\infty z \phi(K, z, T) dz}{\int_0^\infty \phi(K, z, T) dz}.$$

As ϕ is still intractable for small z and computing $z^\beta p(x, z, t)$ is not numerically feasible, we perform an integration by parts. The trade-off is a loss of accuracy due to the computation of $\frac{\partial p}{\partial z}$.

Let us consider the first case above where we are provided with a payoff function f . Using the change of variable solution p , and integrating by parts, we get

$$\begin{aligned} \int_0^\infty f(x) \int_0^\infty \phi(x, z, T) dz dx &= \int_0^\infty f(x) \int_0^\infty z^\beta p(x, z, T) dz dx, \\ &= - \int_0^\infty f(x) \int_0^\infty \frac{z^{\beta+1}}{\beta+1} \frac{\partial p(x, z, T)}{\partial z} dz dx, \end{aligned}$$

since $\beta + 1 > 0$ and $\lim_{z \rightarrow 0} z^{\beta+1} p = \lim_{z \rightarrow 0} z^{\beta+1} p = 0$ due to p being finite. Then we can compute the two quantities above very efficiently using the numerically stable solution p as follows:

$$\mathbb{E}[D_T f(S_T)] = -D_T \int_0^\infty f(x) \int_0^\infty \frac{z^{\beta+1}}{\beta+1} \frac{\partial p(x, z, T)}{\partial z} dz dx \quad (4.19)$$

and

$$\mathbb{E}[V_T | S_T = K] = -(\beta + 1) \frac{\int_0^\infty z^{\beta+1} p(K, z, T) dz}{\int_0^\infty z^{\beta+1} \frac{\partial p(K, z, T)}{\partial z} dz}. \quad (4.20)$$

Finite element method with a two-step BDF time scheme

To the best of our knowledge, most of the literature on the subject (see for instance [10, 38, 52]) relies on ADI schemes to solve the Fokker-Planck PDE in Theorem 5. In this work, we combine a finite element solution with a *Backward Differentiation Formula* (BDF) scheme in time. Thanks to the weak form of the problem, this will allow us to incorporate the Dirac delta initial condition exactly (without using a limiting density function or other smoothing procedure). Furthermore, since Crank-Nicholson time-stepping or ADI schemes can give rise to instabilities (see [35, 51]), we use a two-step implicit BDF scheme instead to get the same second-order convergence in time while preserving stability (we refer to [20] for some financial applications of BDF schemes).

Equation (4.18) can be written as

$$\begin{aligned} \frac{\partial \phi}{\partial t} - z \nabla \cdot \mathbf{u} + \nabla \cdot (\mathbf{b}_1 - (\beta + 1) \mathbf{b}_2) + \mu_t p &= 0, \\ \mathbf{u} &= \frac{1}{2} \begin{bmatrix} \frac{\partial x^2 \alpha^2(x, t) p}{\partial x} + \frac{\partial \rho \xi x \alpha(x, t) p}{\partial z} \\ \frac{\partial \xi^2 p}{\partial z} + \frac{\partial \rho \xi x \alpha(x, t) p}{\partial x} \end{bmatrix}, \quad \mathbf{b}_1 = \begin{bmatrix} (\bar{r}^d(t) - \bar{r}^f(t))x \\ \kappa(\theta - z) \end{bmatrix}, \quad \mathbf{b}_2 = \begin{bmatrix} \rho \xi x \alpha(x, t) p \\ \xi^2 p \end{bmatrix}, \quad \mu_t = -\beta \kappa. \end{aligned}$$

We denote by $\partial\Omega$ the boundary of Ω , by Γ_D the subset with Dirichlet boundary conditions and by Γ_R the subset with Robin boundary condition, so that $\partial\Omega = \Gamma_D \cup \Gamma_R$ where $\Gamma_R = \{(x, z) \in \partial\Omega : z = 0\}$. We derive a weak formulation in the usual way (see, e.g., [36]), i.e., we multiply the PDE by a test function $v \in H^1(\Omega)$, integrate over Ω , using the divergence theorem and boundary conditions, to obtain the weak form of (4.21),

$$\int_{\Omega} \frac{\partial p}{\partial t} v \, d\Omega + a(p, v) = 0,$$

with the bi-linear form

$$\begin{aligned} a(p, v) &= \int_{\Omega} \mathbf{u} \cdot \nabla z v \, d\Omega + \int_{\Omega} \nabla \cdot (\mathbf{b}_1 - (\beta + 1) \mathbf{b}_2) v \, d\Omega + \int_{\Omega} \mu_t p v \, d\Omega \\ &\quad - \int_{\Gamma_R} \left(\kappa z p + \frac{1}{2} z \frac{\partial \rho \xi x \alpha(x, t) p}{\partial x} \right) v \, d\Gamma_R, \end{aligned}$$

where the last term contains the new boundary condition.

The time scheme will be handled by a two-step BDF scheme. Let us define a uniform time discretisation $m \in \llbracket 0, M \rrbracket$ with $t_m = m\Delta_t$. We denote $p_m = p(x, z, t_m)$, in which case the BDF scheme can be written as

$$\int_{\Omega} \left(p_{m+2} - \frac{4}{3} p_{m+1} + \frac{1}{3} p_m \right) v \, d\Omega + \frac{2}{3} \Delta_t a(p_{m+2}, v) = 0,$$

where the first time step is divided into two standard fully implicit time steps (see [21]). At $t = 0$, the Dirac delta initial condition is handled easily via

$$\begin{aligned} \int_{\Omega} p_0 v \, d\Omega &= \int_{\Omega} z^{-\beta} \delta(x - S_0, z - v_0) v \, d\Omega \\ &= v_0^{-\beta} v(S_0, v_0). \end{aligned}$$

Then, for the first time step, we need to solve for p_1 such that

$$\int_{\Omega} p_1 v \, d\Omega + \Delta_t a(p_1, v) = \int_{\Omega} p_0 v \, d\Omega = v_0^{-\beta} v(S_0, v_0),$$

which is equivalent to solving a linear system where the right-hand side vector is $v_0^{-\beta}$ for the source point node (to ensure that the point (S_0, v_0) is on the mesh) and zero otherwise.

The mesh construction for the finite element method is detailed in Appendix A.5.

Calibration algorithm

The calibration of the 2-factor LSV model (2.2) is performed by finding the leverage function α defined in (3.1). In the case of deterministic rates, we know from (3.4) that

$$\alpha(K, T) = \frac{\sigma_{LV}(K, T)}{\sqrt{\mathbb{E}^{\mathbb{Q}}[V_T | S_T = K]}}. \quad (4.21)$$

Hence, we need to compute $\mathbb{E}^{\mathbb{Q}}[V_T | S_T = K]$. The conditional expectation can be written as a function of p , the solution of the dampened PDE (4.18) as given in (4.20). Both integrals can be computed by double adaptive Clenshaw-Curtis quadrature rules to handle singularities properly when T is small (see [23]).

Furthermore, for very small or very large values of K , both the numerator and denominator will be very small. So we define a smooth extrapolation rule as

$$\mathbb{E}^{\mathbb{Q}}[V_T | S_T = K] = \frac{(\beta + 1) \left(\int_0^\infty z^{\beta+1} p(K, z, T) dz + \epsilon (v_0 e^{-\kappa T} + (\theta - v_0) (1 - e^{-\kappa T})) \right)}{\epsilon (\beta + 1) - \int_0^\infty z^{\beta+1} \frac{\partial p(K, z, T)}{\partial v} dz},$$

where we pick $\epsilon = 10^{-14}$ in our numerical tests.

The calibration will be done forward in time. Denote by $(\Delta T(i))_{i \leq N_T}$ the step size and by N_T the number of steps. We use a BDF scheme with constant steps and find that 50 time steps per year and a 80×80 spot-variance finite element mesh give very accurate results.

The leverage function α is defined on a grid of points interpolated with cubic splines in spot and forward-flat in time. There are N_T volatility slices in total such that we denote the i -th slice, i.e., $\alpha(K, T_i)$, by $\alpha_i(K)$. Here, α_i is defined on a grid of N_S points (we find 25 points to provide a sufficiently high numerical accuracy). For a given time slice T_i , the leverage function is defined on the interval $[S_{min}^i, S_{max}^i]$ and is extrapolated flat outside these bounds. This approach allows us to compute $\mathbb{E}^{\mathbb{Q}}[V_T | S_T = K]$ only on the grid points of α , reducing the computational time considerably. Because we need more grid points around the forward value and fewer around S_{min}^i and S_{max}^i , we will use a hyperbolic grid (with $\eta = 0.05$, see Subsection A.5 for more details) refined around the forward value

$$F_i = S_0 e^{\int_0^{T_i} (\bar{r}^d(t) - \bar{r}^f(t)) dt}$$

with

$$S_{min}^i = F_i e^{-\frac{8}{2} \sigma_F(T_i) \sqrt{T_i}}, \quad S_{max}^i = F_i e^{\frac{8}{2} \sigma_F(T_i) \sqrt{T_i}},$$

where $\sigma_F(T_i)$ is the at-the-money forward market volatility for maturity T_i (interpolated linearly in variance). Each of the grid values can be seen as a parameter and we denote them by $(\alpha_{i,j})_{i \leq N_T, j \leq N_S}$, with the associated spot grid points $(s_{i,j})_{i \leq N_T, j \leq N_S}$. We can then write the calibration procedure which we present in Algorithm 1.

In the calibration routine, we use forward-flat interpolation in time for the leverage function to handle the non-linearity of the problem. In between two maturity pillars, where the local volatility is flat, one would expect the exact leverage function to be continuous in time. To account for this, we switch the time interpolation of the leverage function α to linear at the end of the calibration process. The calibrated leverage function α is plotted in Figure 4.2.

Algorithm 1 $\alpha(s, T)$ calibration with the dampened Kolmogorov forward PDE

$$\alpha(s, 0) = \frac{\sigma_{LV}(s, 0)}{\sqrt{v_0}}$$

$T = 0$

for ($i = 1; i \leq N_T; i++$) **do**

solve (4.18) for p on $[T, T + \Delta T]$ with $\alpha(s, [T, T + \Delta T(i)]) = \alpha(s, T)$

for ($j = 1; j \leq N_S; j++$) **do**

$$E_{V_T} = -(\beta + 1) \frac{\int_0^\infty z^{\beta+1} p(s_{i,j}, z, T + \Delta T(i)) dz}{\int_0^\infty z^{\beta+1} \frac{\partial p(s_{i,j}, z, T + \Delta T(i))}{\partial z} dz}$$

$$\alpha_{i,j} = \frac{\sigma_{LV}(s_{i,j}, T + \Delta T(i))}{\sqrt{E_{V_T}}}$$

end for

$T = T + \Delta T(i)$

end for

5 Four-factor Heston-type LSV 2CIR++ model calibration

In this section, we discuss the calibration algorithm for the main model (2.1), which we recall below

$$\begin{cases} \frac{dS_t}{S_t} = (r_t^d - r_t^f) dt + \alpha(S_t, t) \sqrt{V_t} dW_t \\ r_t^d = g_t^d + h^d(t) \\ r_t^f = g_t^f + h^f(t) \\ dg_t^d = \kappa_d(\theta_d - g_t^d) dt + \xi_d \sqrt{g_t^d} dW_t^{g^d} \\ dg_t^f = \left(\kappa_f(\theta_f - g_t^f) - \rho_{Sf} \xi_f \sqrt{g_t^f} \alpha(S_t, t) \sqrt{V_t} \right) dt + \xi_f \sqrt{g_t^f} dW_t^{g^f} \\ dV_t = \kappa(\theta - V_t) dt + \xi \sqrt{V_t} dW_t^V, \end{cases} \quad (5.1)$$

with the correlation structure $d\langle W_t, W_t^V \rangle = \rho dt$, $d\langle W_t, W_t^{g^d} \rangle = \rho_{Sd} dt$, and $d\langle W_t, W_t^{g^f} \rangle = \rho_{Sf} dt$, all other correlations being zero. We will consider the additional model simplification $\rho_{Sd} = \rho_{Sf} = 0$ such that there is no quanto drift adjustment in the foreign short rate dynamics. This allows the use of an extended closed-form formula of [2] presented in Subsection 4.5 when no leverage function appears.

Otherwise, one can extend the approximation formulae of [24] from a 3-factor Heston-CIR to a 4-factor Heston-2CIR++ model, and retrieve the historical correlations between the FX spot and the short-term zero-coupon bonds for ρ_{Sd} and ρ_{Sf} . Alternatively, one can assume $\rho_{Sd} = \rho_{Sf} = 0$ for Subsection 4.5 and perform an initial guess calibration of the stochastic volatility parameters $b_0 = (v_0, \kappa, \theta, \rho, \xi)$. Now, we compute the exact model price C_{MC} at b_0 with non-zero ρ_{sd} and ρ_{sf} using Monte Carlo with the closed-form call price C_A (with $\rho_{Sd} = \rho_{Sf} = 0$) as a control variate, and define the difference between the two prices as $\Delta_0(K, T) = C_{MC} - C_A$. Assuming the absolute values of the two correlations are small, we calibrate the parameter vector $b = (v_0, \kappa, \theta, \rho, \xi)$ on a small interval around b_0 and approximate the call prices by $C_A(b) + \Delta_0$. The calibration algorithm in this section is generic and can be applied to a non-trivial correlation structure.

We assume that both CIR++ processes are calibrated under their own risk-neutral measure and that the Heston-2CIR++ model is calibrated as in Subsection 4.5. The parameters are as follows

$$\begin{cases} v_0 = 0.0094, & \theta = 0.0137, & \kappa = 1.4124, & \rho = -0.1194, & \xi = 0.2988, \\ g_0^d = 0.0001, & \theta_d = 0.5469, & \kappa_d = 0.0837, & \rho_{Sd} = 0, & \xi_d = 0.0274, \\ g_0^f = 0.0001, & \theta_f = 1.1656, & \kappa_f = 0.0110, & \rho_{Sf} = 0, & \xi_f = 0.0370. \end{cases}$$

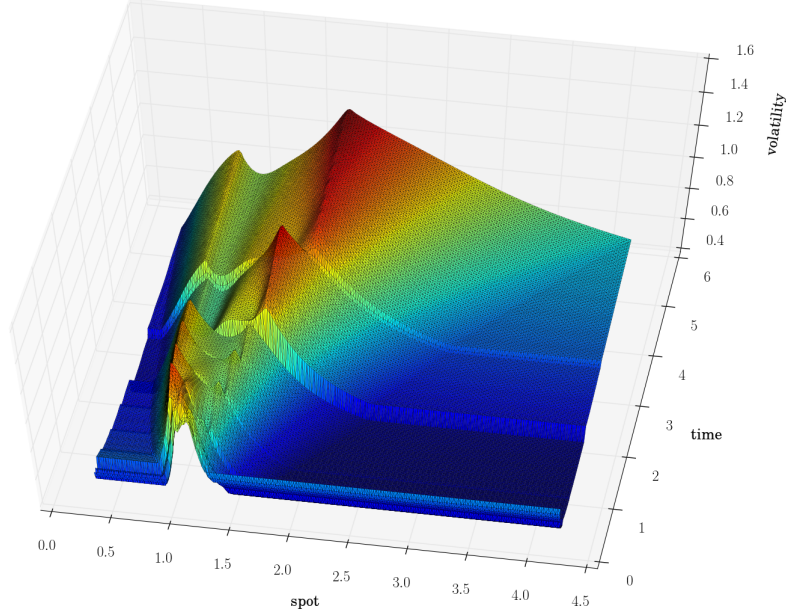


Figure 4.2: Calibrated leverage function α for the 2-factor LSV model (2.2)

In order to run the particle method, we use the Monte Carlo QE-scheme time marching for model (2.1). A full description of the scheme is provided in Appendix A.4.

5.1 Calibration algorithm with the control variate particle method

We recall from (3.1) that the leverage function has to satisfy

$$\alpha^2(K, T) = \frac{\mathbb{E}^{\mathbb{Q}^d} [D_T^d | S_T = K]}{\mathbb{E}^{\mathbb{Q}^d} [D_T^d V_T | S_T = K]} \left(\sigma_{LV}(K, T)^2 + \frac{\mathbb{E}^{\mathbb{Q}^d} [Q_T]}{\frac{1}{2} K^2 \frac{\partial^2 C_{LV}}{\partial K^2}} \right), \quad (5.2)$$

where

$$Q_T = D_T^d \left(r_T^f - \bar{r}^f(T) \right) (S_T - K)^+ - K D_T^d \mathbf{1}_{S_T \geq K} \left[\left(r_T^d - \bar{r}^d(T) \right) - \left(r_T^f - \bar{r}^f(T) \right) \right],$$

which involves conditional expectations as well as standard expectations. Through these expectations, α depends on the distribution of the joint process $X_t = (S_t, V_t, r_t^d, r_t^f)$. The process thus falls in the class of McKean-Vlasov processes [31], and the driving SDE is non-linear. Similarly, the joint density of the process X_t is driven by a non-linear 4-dimensional Kolmogorov forward PDE. The particle method was introduced in [31] and is discussed in detail in [43]; it was applied to LSV model calibration in [25, 26] in order to tackle the non-linearity of the problem.

We consider N -sample paths $(X_t^i)_{i \in \llbracket 1, N \rrbracket} = (S_t^i, V_t^i, r_t^{d,i}, r_t^{f,i})_{i \in \llbracket 1, N \rrbracket}$ of the process X_t . Let Q_T^i be the i -th sample of Q_T . Then α can be estimated by

$$\hat{\alpha}_N(K, T, (X_t^i)_{i \leq N}) = \sqrt{\frac{\sigma_{LV}(K, T)^2}{\hat{p}_N(K, T)} + \frac{\hat{q}_N(K, T)}{\frac{1}{2} \hat{p}_N(K, T) K^2 \frac{\partial^2 C_{LV}}{\partial K^2}}},$$

with

$$\hat{q}_N(s, t) = \frac{1}{N} \sum_{i=1}^N Q_T^i$$

and

$$\hat{p}_N(K, T) = \frac{\sum_{i=1}^N D_T^i V_T^i \delta_N(S_T^i - K)}{\sum_{i=1}^N D_T^i \delta_N(S_T^i - K)},$$

which is an estimator for

$$\frac{\mathbb{E}^{\mathbb{Q}^d} [D_T^d V_T | S_T = K]}{\mathbb{E}^{\mathbb{Q}^d} [D_T^d | S_T = K]},$$

with δ_N a kernel function. The $(4 \times N)$ -dimensional SDE driving the system $(X_t^i)_{i \leq N}$ can be written as

$$\begin{cases} \frac{dS_t^i}{S_t^i} = (r_t^{d,i} - r_t^{f,i}) dt + \hat{\alpha}_N \left(S_t^i, t, (X_t^j)_{j \leq N} \right) \sqrt{V_t^i} dW_t^i \\ r_t^{d,i} = g_t^{d,i} + h^d(t) \\ r_t^{f,i} = g_t^{f,i} + h^f(t) \\ dg_t^{d,i} = \kappa_d (\theta_d - g_t^{d,i}) dt + \xi_d \sqrt{g_t^{d,i}} dW_t^{g^d,i} \\ dg_t^{f,i} = \left(\kappa_f (\theta_f - g_t^{f,i}) - \rho_{Sf} \xi_f \sqrt{g_t^{f,i}} \hat{\alpha}_N \left(S_t^i, t, (X_t^j)_{j \leq N} \right) \sqrt{V_t^i} \right) dt + \xi_f \sqrt{g_t^{f,i}} dW_t^{g^f,i} \\ dV_t^i = \kappa (\theta - V_t^i) dt + \xi \sqrt{V_t^i} dW_t^{V,i}, \end{cases}$$

where $(W_t^i, W_t^{g^d,i}, W_t^{g^f,i}, W_t^{V,i})_{i \in [1, N]}$ are N independent copies of the 4 correlated Brownian motions.

One may notice that all the paths of the $4 \times N$ dimensional process $(X_t^i)_{i \leq N}$ are now entangled due to the dependence of $\hat{\alpha}_N$ in $(X_t^j)_{j \leq N}$. The process can be seen as a system of N interacting particles evolving in a 4 dimensional space, where particle i is defined by its position X_t^i . As pointed out in [25], we will use the term “particle” instead of “path” to account for this behaviour. The density function of the process $(X_t^i)_{i \leq N}$ is now driven by a $(4 \times N)$ -dimensional linear Kolmogorov forward PDE which approximates asymptotically the non-linear 4-dimensional Kolmogorov forward PDE for X_t . Indeed, by the chaos propagation property (we refer to [43] for details) we get that, for any continuous and bounded function f ,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N f(X_T^i) = \int_{(\mathbb{R}^+)^4} f(x) \phi(x) dx,$$

where ϕ is the joint density of (S_t, V_t, r_t^d, r_t^f) , the solution to a non-linear 4D Kolmogorov forward PDE. Therefore, \hat{q}_N and \hat{p}_N converge weakly to the true expectations such that $\hat{\alpha}_N(K, T, (X_T^i)_{i \leq N})$ converges in probability to the solution $\alpha(K, T)$ as $N \rightarrow \infty$.

Remark. The existence and uniqueness of the solution for the SDE are not established theoretically, to the best of our knowledge. From an empirical perspective, [25, 26] show that for very high values of ξ ($\approx 350\%$, that practitioners may require to generate a particular forward skew), the particle method can only calibrate at-the-money strikes. In our case, the ξ ($\approx 30\%$) is calibrated to the market smiles and we are able to reach a high accuracy without any problems.

We will use the 2-factor LSV model with state vector (S_t^{2D}, V_t^{2D}) with the associated leverage function α^{2D} calibrated in Subsection 4.7 as a conditional control variate for \hat{p}_N defined in Subsection 4.3. Additionally, we will use the 1-factor local volatility model with the associated σ_{LV} function calibrated in Subsection 4.6 and keep track of the Dupire forward PDE at time T in line with the Monte Carlo time marching. Indeed, $C_{LV}(K, T)$ and its derivatives appear explicitly in α in (5.2) as well as in the control variates defined in Subsection 4.2.

We denote by N_T the number of time steps and by N the number of particles. The leverage function α is defined on a grid of points and interpolated with cubic splines in spot and forward flat in time. There are $N_T + 1$ volatility “slices” in total such that we denote the m -th time slice, i.e., $\alpha(\cdot, T_m)$, by α_m , represented numerically as splines on a grid of N_S points. While having N_S

too small will lead to accuracy problems, choosing it too large will make the surface rougher due to over-fitting. We find 25-30 points to provide a good trade-off between accuracy and smoothness. For a given time slice T_m , the leverage function is defined on the interval $[S_{min}^m, S_{max}^m]$ and is extrapolated flat outside these bounds. Because we need more grid points around the forward value and less around S_{min}^m and S_{max}^m , we use a hyperbolic grid (with $\eta = 0.05$, see Subsection A.5 for more details) refined around the forward value

$$F_m = S_0 e^{\int_0^{T_m} (\bar{r}^d(t) - \bar{r}^f(t)) dt}$$

with

$$S_{min}^m = F_m e^{-\frac{6}{2} \sigma_F(T_m) \sqrt{T_m}}, \quad S_{max}^m = F_m e^{\frac{6}{2} \sigma_F(T_m) \sqrt{T_m}}.$$

Each of the grid values can be seen as a parameter and we denote them by $(\alpha_{m,j})_{m \leq N_T, j \leq N_S}$ with the associated spot grid values $(s_{m,j})_{m \leq N_T, j \leq N_S}$.

We can now write the calibration algorithm. We denote the 4-factor particle system at time T for the model (2.1) by $(S_T^i, V_T^i, r_T^{d,i}, r_T^{f,i}, D_T^i)_{i \leq N}$, where D_T is an additional state variable for the domestic discount factor. Similarly, we denote the 2-factor particle system at time T for the model (2.2) by $(S_T^{2D,i}, V_T^{2D,i})_{i \leq N}$. We work with an exponential kernel

$$\delta_N(x, T) = \frac{e^{-\frac{1}{2} \left(\frac{x}{h_N(T)} \right)^2}}{h_N(T) \sqrt{2\pi}},$$

with a bandwidth that follows a Silverman-type rule (see [41])

$$h_N(T) = \eta S_0 \sigma_{LV}(S_0, T) \sqrt{\max(T, T_{min})} N^{-\frac{1}{5}},$$

where

$$\eta = 1.5, \quad T = 0.25.$$

The step-by-step calibration is detailed in Algorithm 2.

Our empirical findings suggest that Quasi-Monte Carlo sampling of the random numbers does not provide significant accuracy gains. Also, in order to speed up the computation of the sums involving kernel functions such as

$$\sum_{i=1}^N V_T^{2D,i} \delta_N(S_T^{2D,i} - K),$$

it is advised (see [25]) to sort the particle state vector by spot value and select only the relevant particles that fall inside an interval $[K - \Delta K, K + \Delta K]$, where we choose

$$\Delta K = \sqrt{-2h_N^2(T) \ln(\epsilon \sqrt{2\pi})},$$

with $\epsilon = 10^{-5}$.

5.2 Calibration results

In this subsection, we discuss the results of our calibration and the relative performance improvement with the use of our control variates for the data described in Subsection 4.1. First we compare the benefits of incorporating the control variates (CV) as opposed to the plain particle method (No-CV) as a function of the number of particles. Second, we compare the fit of the 4-factor hybrid LSV model (2.1) with the pure 4-factor hybrid SV model with no leverage function. The error measure we use is the absolute error in volatility (in % units). For instance, a maximum error of 0.03% for a 20% market volatility signifies that the calibrated volatility can be $20.00\% \pm 0.03\%$ in the worst case scenario.

Algorithm 2 $\alpha(s, T)$ Calibration with control variate particle method

$$\alpha(s, T_1 = 0) = \frac{\sigma_{LV}(s, 0)}{\sqrt{v_0}}$$

for ($m = 1; m \leq N_T; m++$) **do**

generate $(Z, Z_v, Z_d, Z_f)_{i \leq N}$ and $(U)_{i \leq N}$, i.e. $4 \times N$ independent draws from $\mathcal{N}(0, 1)$ and N draws from $\mathcal{U}([0, 1])$, respectively

evolve the 4-factor particle system from T_m to T_{m+1} with QE -Scheme (A.15) where $\alpha(s, [T_m, T_{m+1}]) = \alpha(s, T_m)$

evolve the 2-factor particle system from T_m to T_{m+1} with QE -Scheme (A.15) with pre-computed α^{2D} and using $(Z, Z_v, U)_{i \leq N}$

solve the Dupire forward PDE (4.14) from T_m to T_{m+1} for C_{LV} , $\frac{\partial C_{LV}}{\partial K}$ and $\frac{\partial^2 C_{LV}}{\partial K^2}$

set $T = T_{m+1}$

for ($j = 1; j \leq N_S; j++$) **do**

set $K = s_{m+1, j}$

compute as in (4.21)

$$p^{2D} = \mathbb{E}^{\mathbb{Q}^d} [V_T^{2D} | S_T^{2D} = K] = \left(\frac{\sigma_{LV}(K, T)}{\alpha^{2D}(K, T)} \right)^2$$

compute as in (4.3)

$$\hat{p}_N^{2D}(K, T) = \frac{\sum_{i=1}^N V_T^{2D, i} \delta_N(S_T^{2D, i} - K)}{\sum_{i=1}^N \delta_N(S_T^{2D, i} - K)}$$

compute $\hat{\lambda}_N$ as in (4.5) and

$$p_N^*(K, T) = \frac{\sum_{i=1}^N D_T^i V_T^i \delta_N(S_T^i - K)}{\sum_{i=1}^N D_T^i \delta_N(S_T^i - K)} + \hat{\lambda}_N (\hat{p}_N^{2D}(K, T) - p^{2D}(K, T))$$

as in (4.4)

compute as in (4.2)

$$\gamma_N^* = (\gamma_{1, N}^* - K \gamma_{2, N}^*)$$

with

$$\gamma_{1, N}^* = \hat{\gamma}_{1, N} + \lambda_1 (\hat{\mu}_{1, N} - C_{LV}(K, T)) + \eta_1 \left(\hat{\zeta}_{1, N} - \mathbb{E}^{\mathbb{Q}^d} \left[\left(r_T^f - \bar{r}^f(T) \right) \right] \right)$$

$$\gamma_{2, N}^* = \hat{\gamma}_{2, N} + \lambda_2 \left(\hat{\mu}_{2, N} + \frac{\partial C_{LV}(K, T)}{\partial K} \right) + \eta_2 \left(\hat{\zeta}_{2, N} - \mathbb{E}^{\mathbb{Q}^d} \left[\left(r_T^d - \bar{r}^d(T) \right) - \left(r_T^f - \bar{r}^f(T) \right) \right] \right)$$

compute

$$\alpha_{m+1, j} = \sqrt{\frac{1}{p_N^*(K, T)} \left(\sigma_{LV}(K, T)^2 + \frac{\gamma_N^*(K, T)}{\frac{1}{2} K^2 \frac{\partial^2 C_{LV}(K, T)}{\partial K^2}} \right)}$$

end for

end for

We use the calibration routine described in Subsection 5.1 with and without control variates for 800, 4 000, 20 000, 100 000, 500 000 and 2 500 000 particles. Additionally, we set the number of time steps to 250 per year (such that, on average, there is one time step per open day). Both the average and maximum error are computed on all quoted deltas and maturities defined in Subsection 4.1. The results are presented in Figure 5.1.

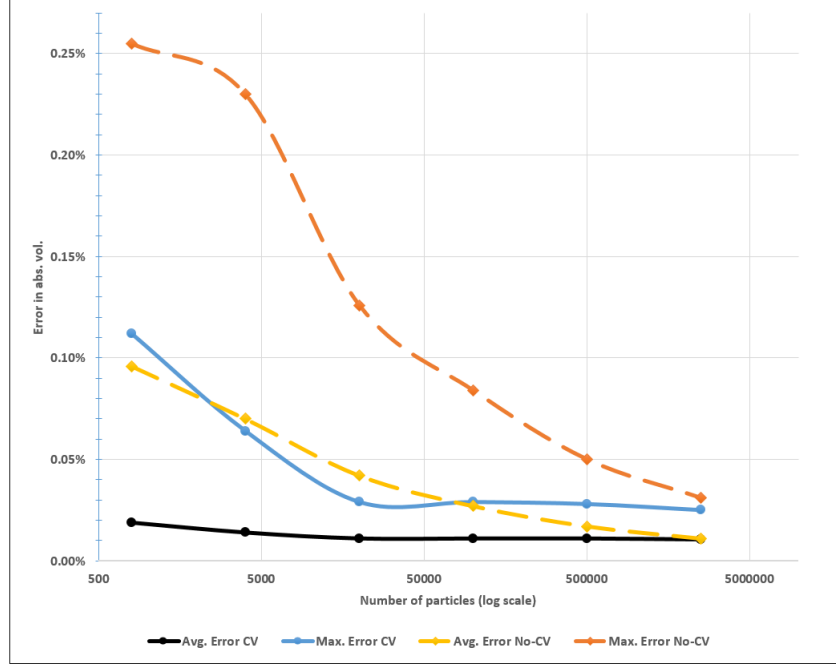


Figure 5.1: Error convergence with and without control variates. Error is computed on all quoted deltas and maturities (log-scale)

We infer from the data in Figure 5.1 that the use of control variates greatly improves the general calibration routine. On the one hand, convergence¹ in both the maximum error (0.029%) and in the average error (0.011%) is reached for 20 000 particles when using control variates. On the other hand, the plain particle method without the variance reduction can only reach the same accuracy with 2 500 000 particles. In order to emphasise this fact, we calibrate the 4-factor model with 4 000 particles with and without control variates. We then plot, in Figure 5.2, the model implied volatility slices 3M, 1Y, 2Y, 5Y.

Hence, from Figure 5.1, we infer that the control variate particle method converges approximately 125 times faster than the plain particle method (without control variates). Most practitioners are familiar with the computational cost for the calibration of the 2-factor LSV model (2.2). One only needs to add the diffusion of 20 000 particles to get a very accurate calibration to vanilla options for the 4-factor LSV model (2.1). This allows us to calibrate a 4-factor model almost as fast as the 2-factor version. Moreover, from careful data analysis of Figure 5.1, we notice a clear convergence rate of 0.3 in the number of particles for the plain particle method, for both the average and maximum error. The addition of the control variates preserve the convergence rate but reduces the absolute size of the error significantly. We note that the error can be further reduced by increasing the number of time steps per year. However, we decided to work with 250 time steps (per year) as it already yields a very accurate calibration at a reasonable computational cost.

Part of the variance reduction is due to the conditional control variate described in Subsection 4.3 which can provide very good results for short-term horizons. In order to understand this, we

¹i.e., the error from here on is dominated by other sources, such as the time discretisation error

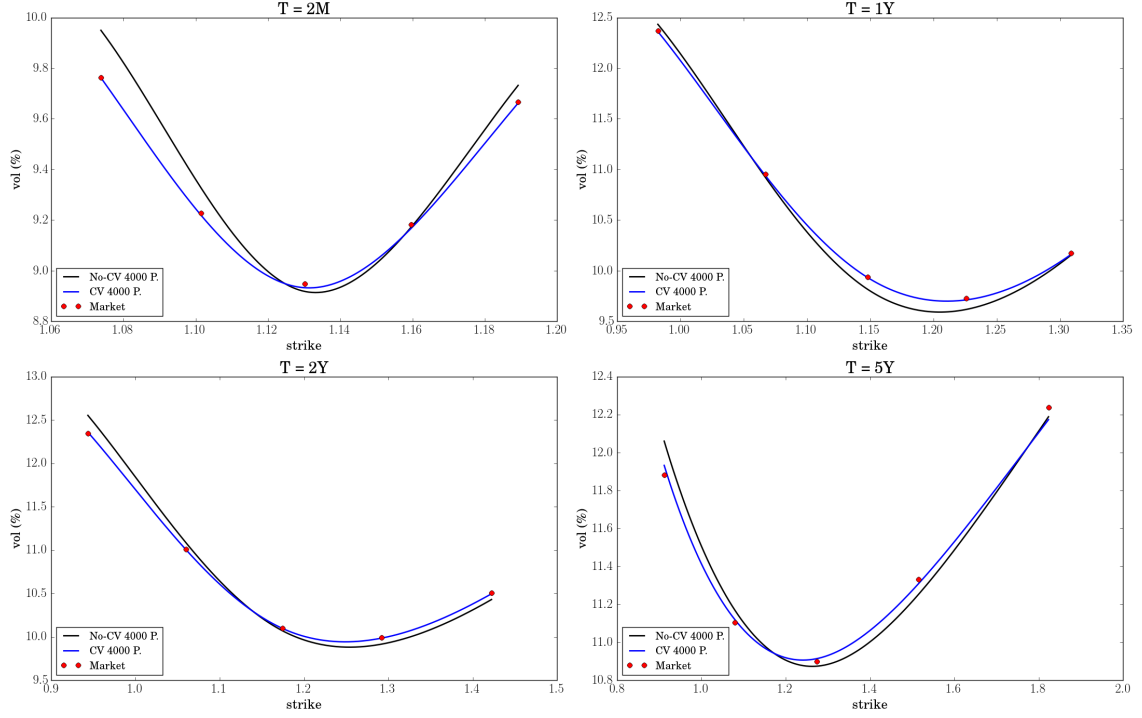


Figure 5.2: Calibration fit with 4 000 particles for the 4-factor LSV model with and without control variates

recall the variance reduction factor in (4.6),

$$\frac{1}{1 - \text{Corr}(m, m^{2D})^2}.$$

We plot the quantity as a function of time with 500 000 particles in Figure 5.3. We notice a trend-line in $1 + \frac{C}{T^2}$, for a given constant C , and thus a very good variance reduction for short-term options calibration. For longer term, this still yields a good variance reduction factor of 2.7 for the 3Y maturity and 1.44 for the 5Y pillar. In addition, the other two control variates, presented in Subsection 4.2, help control the short- to long-term behaviour as well. The variance reduction factor for $\gamma_{1,N}^*$ and $\gamma_{2,N}^*$ seems to reach a steady state in time around 1.6 and 6.5, respectively. We plot both variance reduction factors in Figure 5.3.

As a final observation on the calibration fit, we present and compare the accuracy of the 4-factor hybrid LSV model with the 4-factor Hybrid SV (Heston–2CIR++ model without leverage function that was calibrated in Subsection 4.5. We use 500 000 particles for that purpose and the results are presented in Figure 5.4 as well as in Table 3. We included a short-term pillar (3M), two mid-term pillars (1Y, 2Y) as well as one long-term pillar (5Y). The re-pricing is performed by Monte Carlo with 500 time steps per year and 5 000 000 paths. The average calibration error is 0.011%, which means that for a market volatility of 20%, the calibrated volatility is $20\% \pm 0.011\%$, on average.

The associated leverage function is plotted in Figure 5.5.

We notice that the short- to mid-term leverage function is smooth and well behaved. However, it becomes less stable around 5Y for out-of-the-money spot values. This is expected since for short- and mid-term pillars the variance reduction obtained with the 2-factor LSV model is significant. Longer term and out-of-the-money spot values are harder to estimate accurately because of the term

$$\frac{\mathbb{E}^{\mathbb{Q}^d} \left[D_T^d \left(r_T^f - \bar{r}^f(T) \right) (S_T - K)^+ \right] - K \mathbb{E}^{\mathbb{Q}^d} \left[D_T^d \mathbf{1}_{S_T \geq K} \left[\left(r_T^d - \bar{r}^d(T) \right) - \left(r_T^f - \bar{r}^f(T) \right) \right] \right]}{\frac{1}{2} K^2 \frac{\partial^2 C_{LV}}{\partial K^2}}$$

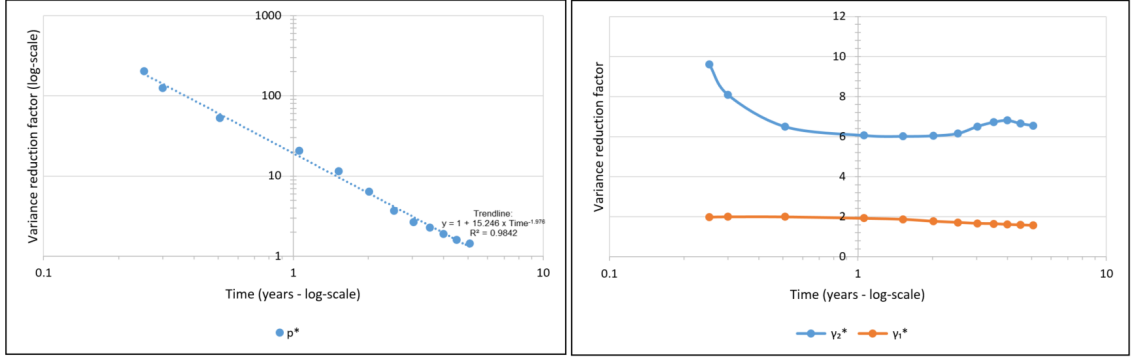


Figure 5.3: Variance reduction factor as a function of time for the conditional expectation estimator p_N^* (left, using a log-log scale) and the standard expectations estimators $\gamma_{1,N}^*$ and $\gamma_{2,N}^*$ (right, using a log scale)

	10D Put		25D Put		50D		25D Call		10D Call		Avg.	
T	LSV	SV	LSV	SV	LSV	SV	LSV	SV	LSV	SV	LSV	SV
3W	0.01	1.14	0.01	1.19	0.02	1.08	0.02	0.69	0.02	0.28	0.02	0.88
1M	0.02	1.33	0.01	1.29	0.01	1.14	0.00	0.78	0.00	0.47	0.01	1.00
2M	0.00	0.72	0.01	0.61	0.01	0.48	0.01	0.29	0.00	0.20	0.01	0.46
3M	0.01	0.12	0.02	0.03	0.02	0.09	0.02	0.08	0.01	0.03	0.02	0.07
6M	0.03	0.68	0.03	0.63	0.02	0.36	0.02	0.05	0.01	0.40	0.02	0.42
1Y	0.01	0.60	0.01	0.56	0.01	0.29	0.01	0.14	0.01	0.48	0.01	0.41
1Y6M	0.01	0.48	0.00	0.40	0.01	0.13	0.01	0.22	0.01	0.49	0.01	0.34
2Y	0.01	0.37	0.00	0.26	0.01	0.03	0.01	0.30	0.01	0.50	0.01	0.29
3Y	0.02	0.20	0.00	0.18	0.00	0.10	0.00	0.01	0.01	0.07	0.01	0.11
5Y	0.01	0.22	0.00	0.25	0.01	0.07	0.00	0.27	0.02	0.61	0.01	0.29

Table 3: Calibration error in absolute volatility (% unit) for the 4-factor LSV model vs 4-factor Heston–2CIR++ model

that appears in $\alpha(K, T)$. Indeed, the computation of the two expectations is less accurate out-of-the-money (since the relative variances of the two payoffs are bigger) and a small error is emphasised by the very small denominator. Thankfully, the use of control variates from 4.2 combined with a good estimation of the denominator (as explained in Subsection 4.6) allows to dampen the oscillations. Furthermore, the calibration fit is not very sensitive to far out-of-(or deep in-)the-money values of the leverage function.

5.3 Stress scenario

The last calibration result we present is for an extremely stressed set of rates parameters. This will serve as a robustness test of the control-variate particle method for very volatile short rate processes. In order to perform the test, we multiply the calibrated ξ_d and ξ_f by 20 as well as divide θ_d and θ_f by 20. This leads to strongly violated Feller conditions and very high volatilities for the two CIR++ rate processes. The stress scenario parameter values are

$$\begin{cases} v_0 = 0.0094, & \theta = 0.0137, & \kappa = 1.4124, & \rho = -0.1194, & \xi = 0.2988, \\ g_0^d = 0.0001, & \theta_d = \mathbf{0.0273}, & \kappa_d = 0.0837, & \rho_{Sd} = 0, & \xi_d = \mathbf{0.5480}, \\ g_0^f = 0.0001, & \theta_f = \mathbf{0.0582}, & \kappa_f = 0.0110, & \rho_{Sf} = 0, & \xi_f = \mathbf{0.7400}. \end{cases}$$

We emphasise that in practice, the volatility parameters $\xi_{d/f}$ are rarely above 0.06. One can refer to [6] for more details. A calibration summary for the average error in absolute volatility is displayed

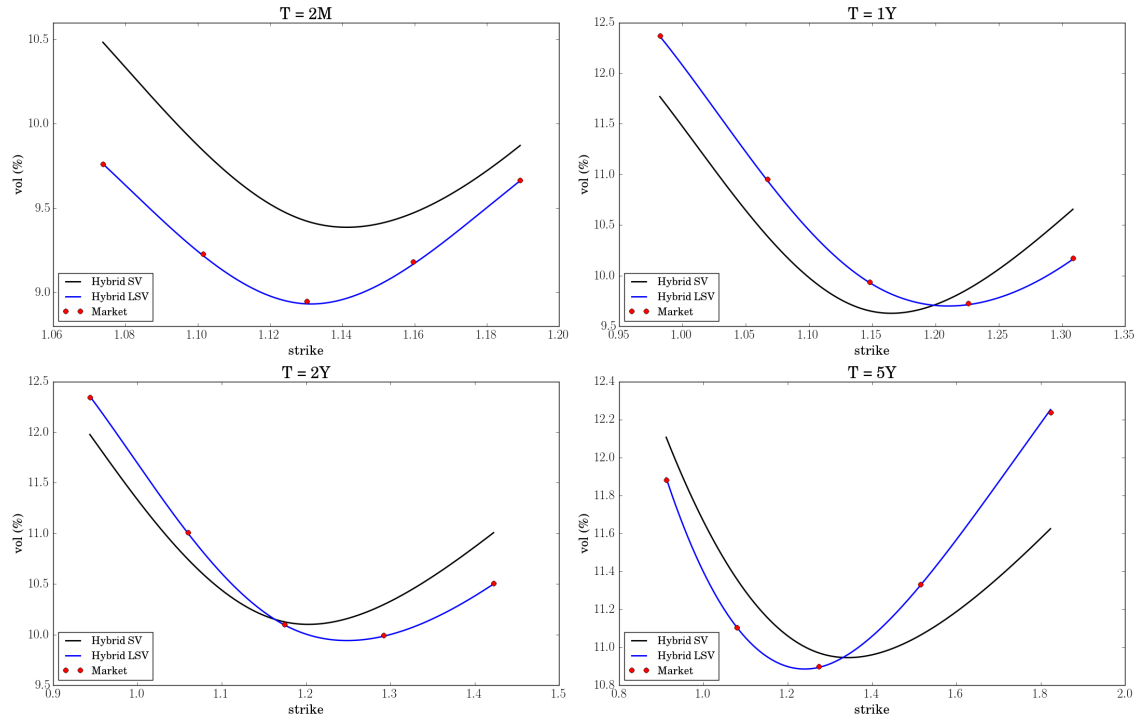


Figure 5.4: Calibration fit with 500 000 particles for the 4-factor LSV model vs 4-factor SV model

in Table 4. On this stress scenario, the calibration via control-variate particle method reaches an error of 0.0165% for 20 000 particles, whereas 2 500 000 particles are required with the plain particle method (without control variates) to reach the same accuracy. This is in line with Figure 5.1 for our previous observations with realistic short rate parameters calibrated to the market. Hence, the control-variate particle method shows a consistent improvement over the plain particle method even under stress scenarios. The conditional control variate alone yields a good variance reduction factor of almost 2 for the last maturity (5Y) .

N	4 000	20 000	100 000	500 000	2 500 000
Plain particle method	0.0855%	0.0440%	0.0320%	0.0255%	0.0160%
Control variate particle method	0.0312%	0.0165%	0.0133%	0.0133%	0.0133%

Table 4: Average error in absolute volatility (% unit) for a high-volatility stress scenario on the rate processes. N is the number of particles.

6 Conclusion

In this paper, we have provided a new and numerically effective method to calibrate a 4-factor LSV model to vanilla options. According to our numerical results and based on real world market data, we managed to achieve an approximate 125-fold speed-up for the calibration using control variates, as compared to the plain particle method. We have shown that a high accuracy can be obtained with as few as 20 000 particles (with a maximum error in absolute volatility of 0.03%), and we were able to get a good fit with only 4 000 particles (with a maximum error in absolute volatility of 0.06%). Using the calibrated leverage function from this paper, we showed in [14] that the addition of stochastic rates may have a significant impact on structured products, even more so when barrier features and coupon detachments are combined for longer-dated contracts. Stochastic

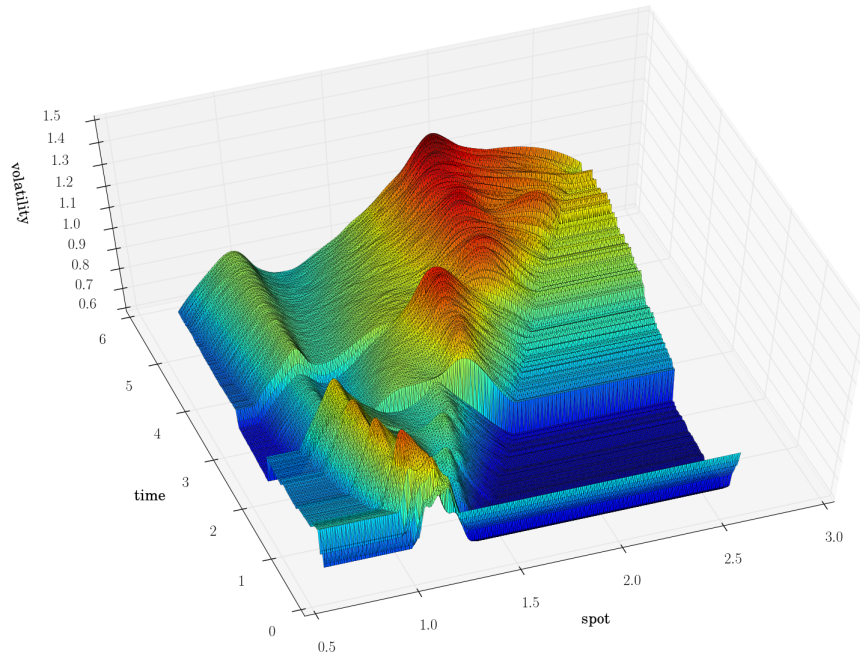


Figure 5.5: Calibrated leverage function α for the 4-factor LSV model (2.1)

rates become necessary in the modelling if one wishes to price hybrid products where the rates appear explicitly (for instance, a spread option on the FX performance and the Libor rate). One could use a second factor in the CIR++ processes to improve the fit to caps and use the method presented in this article to calibrate the leverage function.

References

- [1] M. ABRAMOWITZ, *Handbook of Mathematical Functions, with Formulas, Graphs, and Mathematical Tables*, Dover Publications, 1974.
- [2] R. AHLIP AND M. RUTKOWSKI, *Pricing of foreign exchange options under the Heston stochastic volatility model and CIR interest rates*, Quantitative Finance, 13 (2013), pp. 955–966.
- [3] L. ANDERSEN, *Simple and efficient simulation of the Heston stochastic volatility model*, The Journal of Computational Finance, 11 (2008), pp. 1–42.
- [4] L. ANDERSEN AND V. PITERBARG, *Moment explosions in stochastic volatility models*, Finance and Stochastics, 11 (2007), pp. 29–50.
- [5] D. BRIGO AND F. MERCURIO, *A deterministic-shift extension of analytically-tractable and time-homogeneous short-rate models*, Finance and Stochastics, 5 (2001), pp. 369–387.
- [6] D. BRIGO AND F. MERCURIO, *Interest Rate Models: Theory and Practice*, Springer, 2006.
- [7] M. BROADIE AND O. KAYA, *Exact simulation of stochastic volatility and other affine jump diffusion processes*, Operations Research, 54 (2006), pp. 217–231.
- [8] R. CARMONA AND S. NADTOCHIY, *Local volatility dynamic models*, Finance and Stochastics, 13 (2009), pp. 1–48.
- [9] P. CARR, *Implied vol constraints*, tech. report, Bloomberg Working Paper, 2004. Available at www.javaquant.net/papers/impvolconstrs3.pdf.
- [10] I. J. CLARK, *Foreign Exchange Option Pricing: A Practitioner’s Guide*, John Wiley & Sons, 2010.
- [11] T. COLEMAN AND Y. LI, *On the convergence of interior-reflective Newton methods for nonlinear minimization subject to bounds*, Mathematical Programming, 67 (1994), pp. 189–224.

- [12] T. F. COLEMAN AND Y. LI, *An interior trust region approach for nonlinear minimization subject to bounds*, SIAM Journal on Optimization, 6 (1996), pp. 418–445.
- [13] J. COX, J. INGERSOLL, AND S. ROSS, *A theory of the term structure of interest rates*, Econometrica, 53 (1985), pp. 385–407.
- [14] A. COZMA, M. MARIAPRAGASSAM, AND C. REISINGER, *Convergence of an Euler scheme for a hybrid stochastic-local volatility model with stochastic rates in foreign exchange markets*. arxiv preprint, arXiv:1501.06084.
- [15] S. CRÉPEY, *Tikhonov regularization*, Encyclopedia of Quantitative Finance, (2010), pp. 1807–1812.
- [16] G. DEELSTRA AND G. RAYEE, *Local volatility pricing models for long-dated FX derivatives*, Applied Mathematical Finance, 20 (2013), pp. 380–402.
- [17] P. DIERCKX, *An algorithm for surface-fitting with spline functions*, IMA Journal of Numerical Analysis, 1 (1981), pp. 267–283.
- [18] B. DUPIRE, *A unified theory of volatility*, in Derivatives Pricing: The Classic Collection, Risk Books, 2004.
- [19] M. R. FENGLER, *Arbitrage-free smoothing of the implied volatility surface*, Quantitative Finance, 9 (2009), pp. 417–428.
- [20] F. L. FLOCH, *TR-BDF2 for stable American option pricing*, Journal of Computational Finance, 17 (2014).
- [21] M. GILES AND R. CARTER, *Convergence analysis of Crank-Nicolson and Rannacher time-marching*, Journal of Computational Finance, 9 (2006), pp. 89–112.
- [22] P. GLASSERMAN, *Monte Carlo Methods in Financial Engineering*, vol. 53 of Stochastic Modelling and Applied Probability, Springer, 2003.
- [23] B. GOUGH, *GNU Scientific Library Reference Manual - Third Edition*, Network Theory Ltd., 3rd ed., 2009.
- [24] L. A. GRZELAK AND C. W. OOSTERLEE, *On the Heston model with stochastic interest rates*, SIAM Journal on Financial Mathematics, 2 (2011), pp. 255–286.
- [25] J. GUYON AND P. HENRY-LABORDERE, *The smile calibration problem solved*. SSRN.1885032, 2011.
- [26] J. GUYON AND P. HENRY-LABORDERE, *Nonlinear Option Pricing*, Chapman and Hall/CRC Financial Mathematics, Chapman and Hall/CRC, 2013.
- [27] I. GYÖNGY, *Mimicking the one-dimensional marginal distributions of processes having an Itô differential*, Probability Theory and Related Fields, 71 (1986), pp. 501–516.
- [28] T. R. HURD AND A. KUZNETSOV, *Explicit formulas for Laplace transforms of stochastic integrals*, Markov Processes and Related Fields, 14 (2008), pp. 277–290.
- [29] P. JÄCKEL, *Let's be rational*, Wilmott, 2015 (2015), pp. 40–53.
- [30] V. LUCIC, *Boundary conditions for computing densities in hybrid models via PDE methods*, Stochastics: An International Journal of Probability and Stochastic Processes, 84 (2012), pp. 705–718.
- [31] H. P. MCKEAN, *A class of Markov processes associated with nonlinear parabolic equations*, Proceedings of the National Academy of Sciences of the United States of America, 56 (1966), pp. 1907–1911.
- [32] MUREX ANALYTICS, *Local volatility model*, Murex, (2007). Internal.
- [33] R. T. O.C. ZIENKIEWICZ AND J. ZHU, *The Finite Element Method Set. Its Basis and Fundamentals*, Butterworth-Heinemann, 6 ed., 2005.
- [34] B. ØKSENDAL, *Stochastic Differential Equations: An Introduction with Applications (Universitext)*, Springer, Jan 2014.
- [35] D. M. POOLEY, K. R. VETZAL, AND P. A. FORSYTH, *Convergence remedies for non-smooth payoffs in option pricing*, Journal of Computational Finance, 6 (2003), pp. 25–40.
- [36] A. QUARTERONI AND A. VALLI, *Numerical approximation of partial differential equations*, vol. 23, Springer Science & Business Media, 2008.
- [37] A. REHAI, *The hybrid most likely path*, Risk Magazine, April (2006).
- [38] Y. REN, D. MADAN, AND M. QIAN QIAN, *Calibrating and pricing with embedded local volatility models*, Risk Magazine, September (2007).

- [39] C. L. C. G. ROGERS AND D. WILLIAMS, *Diffusions, Markov processes, and Martingales. Vol. 2*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 2000.
- [40] R. SCHÖBEL AND J. ZHU, *Stochastic volatility with an Ornstein–Uhlenbeck process: an extension*, European Finance Review, 3 (1999), pp. 23–46.
- [41] B. W. SILVERMAN, *Density estimation for statistics and data analysis*, Biometrical Journal, 30 (1988), pp. 876–877.
- [42] R. STORN AND K. PRICE, *Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces*, Journal of Global Optimization, 11 (1997), pp. 341–359.
- [43] A.-S. SZNITMAN, *Ecole d’Eté de Probabilités de Saint-Flour XIX*, Springer, 1991, ch. Topics in propagation of chaos, pp. 165–251.
- [44] Y. TIAN, Z. ZHU, G. LEE, F. KLEBANER, AND K. HAMZA, *Calibrating and pricing with a stochastic-local volatility model*, Journal of Derivatives, 22 (2015), pp. 21–39.
- [45] L. TUR, *Local volatility calibration with fixed-point algorithm*, GDF Suez Trading, (2014). Internal.
- [46] A. VAN DER STOEP, L. A. GRZELAK, AND C. W. OOSTERLEE, *The Heston stochastic-local volatility model: efficient Monte Carlo simulation*, International Journal of Theoretical and Applied Finance, 17 (2014), pp. 1–30.
- [47] A. W. VAN DER STOEP, L. A. GRZELAK, AND C. W. OOSTERLEE, *A novel Monte Carlo approach to hybrid local volatility models*. preprint, SSRN <http://ssrn.com/abstract=2766990>, 2016.
- [48] A. VAN HAASTRECHT, R. LORD, A. PELSSER, AND D. SCHRAGER, *Pricing long-maturity equity and FX derivatives with stochastic interest rates and stochastic volatility*, Insurance: Mathematics and Economics, 45 (2009), pp. 436–448.
- [49] A. VAN HAASTRECHT AND A. PELSSER, *Generic pricing of FX, inflation and stock options under stochastic interest rates and stochastic volatility*, Quantitative Finance, 11 (2011), pp. 665–691.
- [50] R. WHITE, *Numerical solution to PDEs with financial applications*, OpenGamma Quantitative Research, (2013). <https://developers.opengamma.com/quantitative-research/numerical-solutions-to-pdes-with-financial-applications-opengamma.pdf>.
- [51] M. WYNS, *Convergence analysis of the Modified Craig-Sneyd scheme for two-dimensional convection-diffusion equations with nonsmooth initial data*, arXiv preprint arXiv:1508.04296, (2015).
- [52] M. WYNS AND J. D. TOIT, *A finite volume - alternating direction implicit approach for the calibration of stochastic local volatility models*, arXiv preprint arXiv:1611.02961, (2016).

A Additional results

A.1 Derivation of Lemmas 2 and 3

A.1.1 Proof of Lemma 2

We first state a necessary auxiliary result, which is an adaptation of a result in [39], where the integrand D in the local time integral is 1.

Proposition 7. *On a filtered probability space $(\mathcal{X}, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{Q}^d)$, let D and X be two \mathcal{F}_t -adapted continuous semi-martingales, with D positive and integrable, $(l_t^a)_{t \geq 0}$ the local time of X at level a and, for all $n > 0$,*

$$\delta_n^a(x) = \begin{cases} 0, & |x - a| > \frac{1}{n}, \\ \frac{n}{2}, & |x - a| \leq \frac{1}{n}. \end{cases} \quad (\text{A.1})$$

Then, for any horizon time $T > 0$ and $0 < t \leq T$, $\int_0^t D_s \delta_n^a(X_s) d\langle X \rangle_s$ converges almost surely, and uniformly in time, to $\int_0^t D_s d_s l^a$.

Proof. We closely follow Section 45 of [39], but use a more concrete expression of the regularisation function. The local time at level a is defined as a continuous adapted increasing process such that

$$|X_t - a| - |X_0 - a| = \int_0^t \operatorname{sgn}(X_s - a) dX_s + l_t^a, \quad (\text{A.2})$$

with $\text{sgn}(x) = -1$ for $x \leq 0$ and $\text{sgn}(x) = 1$ for $x > 0$. We define a sequence of functions $(f_n)_{n \geq 0}$ for all $x \in \mathbb{R}$, as in Chapter 4 of [34]

$$f_n(x) = \begin{cases} |x - a|, & |x - a| > \frac{1}{n}, \\ \frac{1}{2} \left(\frac{1}{n} + n(x - a)^2 \right), & |x - a| \leq \frac{1}{n}. \end{cases}$$

Hence, for all $n > 0$, $\frac{1}{2}f_n'' = \delta_n^a$ a.e. We recall from the proof of Tanaka's formula (see [39] for more details) that $\int_0^t \delta_n^a(X_s) d\langle X \rangle_s$ converges almost surely to l_t^a (uniformly in t). The sequence f_n converges uniformly to $x \rightarrow |x - a|$ and f_n' converges point-wise to $\text{sgn}(x - a)$ such that by the Itô-Doeblin formula we can write

$$\begin{aligned} D_t f_n(X_t) - D_0 f_n(X_0) &= \int_0^t f_n(X_s) dD_s + \int_0^t D_s f_n'(X_s) dX_s \\ &+ \frac{1}{2} \int_0^t D_s f_n''(X_s) d\langle X \rangle_s + \int_0^t d\langle D, f_n(X) \rangle_s. \end{aligned} \quad (\text{A.3})$$

We denote

$$C_t^n = \frac{1}{2} \int_0^t D_s f_n''(X_s) d\langle X \rangle_s$$

and, since $f_n''(x) = 0$ for any x such that $|x - a| \geq \frac{1}{n}$, we have

$$\int_0^t \mathbf{1}_{|X_s - a| > \frac{1}{n}} dC_s^n = 0.$$

Also, from the definition of f_n , for all $x \in \mathbb{R}$,

$$\text{sgn}(x - a) - f_n'(x) = \begin{cases} 0, & |x - a| \geq \frac{1}{n}, \\ \text{sgn}(x - a) - n(x - a), & |x - a| < \frac{1}{n}, \end{cases}$$

and then for any given x , and $n > 0$, both $D_t |\text{sgn}(x - a) - f_n'(x)|$ and $D_t ||x - a| - f_n(x)|$ are smaller than D_t which is integrable. Let $X_t = X_0 + M_t + A_t$ be the canonical decomposition of X and $D_t = D_0 + N_t + R_t$ the canonical decomposition of D . Localisation allows us to reduce the problem to the case where M and N are bounded and A and R are of bounded variation. Then,

$$\left\| \int_0^T D_s \left(\text{sgn}(X_s - a) - f_n'(X_s) \right) dM_s \right\|_2^2 = \mathbb{E} \left[\int_0^T \left(D_s \left(\text{sgn}(X_s - a) - f_n'(X_s) \right) \right)^2 d\langle M \rangle_s \right],$$

for which the right-hand-side goes to zero when n goes to infinity. By Doob's L^2 martingale inequality, we can write

$$\left\| \sup_{t \in [0, T]} \left| \int_0^t D_s \left(\text{sgn}(X_s - a) - f_n'(X_s) \right) dM_s \right| \right\|_2 \leq 2 \left\| \int_0^T D_s \left(\text{sgn}(X_s - a) - f_n'(X_s) \right) dM_s \right\|_2,$$

to conclude that

$$\sup_{t \in [0, T]} \left| \int_0^t D_s \left(\text{sgn}(X_s - a) - f_n'(X_s) \right) dM_s \right| \rightarrow 0,$$

in L^2 and in probability. We may then assume that (A.1.1) also holds almost surely, since, we could work with a subsequence on n for which the statement is true instead). Similarly, we have

$$\sup_{t \in [0, T]} \left| \int_0^t (|X_t - a| - f_n(X_s)) dN_s \right| \rightarrow 0 \text{ a.s.}$$

Also,

$$\begin{aligned} \left| \int_0^t D_s \left(\operatorname{sgn}(X_s - a) - f'_n(X_s) \right) dA_s \right| &\leq \int_0^t D_s \left| \left(\operatorname{sgn}(X_s - a) - f'_n(X_s) \right) \right| |dA_s| \\ &\leq \int_0^t D_s |dA_s|, \end{aligned}$$

and

$$\sup_{t \in [0, T]} \left| \int_0^t D_s \left(\operatorname{sgn}(X_s - a) - f'_n(X_s) \right) dA_s \right| \leq \int_0^T D_s |dA_s|.$$

The monotone-convergence theorem allows us to conclude that

$$\sup_{t \in [0, T]} \left| \int_0^t D_s \left(\operatorname{sgn}(X_s - a) - f'_n(X_s) \right) dA_s \right| \rightarrow 0,$$

in L^1 , in probability and almost surely (on passing to a sub-sequence on n if necessary). Hence,

$$\int_0^t D_s f'_n(X_s) dA_s \rightarrow \int_0^t D_s \operatorname{sgn}(X_s) dA_s \quad (\text{A.4})$$

almost surely and uniformly in time. Similarly,

$$\left| \int_0^t (f_n(X_s) - |X_s - a|) dR_s \right| \leq \int_0^t |dR_s|,$$

and we get

$$\int_0^t f_n(X_s) dR_s \rightarrow \int_0^t |X_s - a| dR_s$$

almost surely and uniformly in time. Additionally, we can write

$$\int_0^t d \langle D, f_n(X) \rangle_s = \int_0^t f'_n(X_s) d \langle N, M \rangle_s.$$

From the Kunita-Watanabe inequality,

$$\begin{aligned} \int_0^t \left| d \langle D, (\operatorname{sgn}(X_s - a) - f'_n(X_s)) \rangle_s \right| &\leq \int_0^t \left| (\operatorname{sgn}(X_s - a) - f'_n(X_s)) \right| |d \langle N, M \rangle_s| \\ &\leq \sqrt{\int_0^t |(\operatorname{sgn}(X_s - a) - f'_n(X_s))|^2 d \langle M \rangle_s} \sqrt{\int_0^t d \langle N \rangle_s}. \end{aligned}$$

Since $\langle N \rangle_s$ and $\langle M \rangle_s$ are increasing processes of finite variation, we proceed as in (A.4) and conclude

$$\int_0^t d \langle D, f'_n(X) \rangle_s \rightarrow \int_0^t d \langle D, \operatorname{sgn}(X - a) \rangle_s \text{ a.s.}$$

Hence, from (A.3), C^n converges to a limit ζ almost surely (uniformly in time t). Applying integration by parts to the Tanaka formula (A.2), we can write

$$\begin{aligned} D_t |X_t - a| - D_0 |X_0 - a| &= \int_0^t |X_s - a| dD_s + \int_0^t D_s \operatorname{sgn}(X_s - a) dX_s \\ &+ \int_0^t D_s d_s l^a + \int_0^t d \langle D, |X - a| \rangle_s \end{aligned}$$

and conclude that

$$\zeta_t = \int_0^t D_s d_s l^a.$$

□

We can now prove Lemma 2.

Proof of Lemma 2. From Proposition 7, we know that the following holds almost surely:

$$\int_0^t D_u dl_u^a = \lim_{n \rightarrow \infty} \int_0^t D_u \delta_n^a(X_u) Y_u^2 du.$$

This implies convergence in distribution, so that we can write

$$\mathbb{E}^{\mathbb{Q}^d} \left[\int_0^t D_u dl_u^a \right] = \lim_{n \rightarrow \infty} \mathbb{E}^{\mathbb{Q}^d} \left[\int_0^t D_u \delta_n^a(X_u) Y_u^2 du \right],$$

and, by the stochastic Fubini theorem, we get

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}^d} \left[\int_0^t D_u dl_u^a \right] &= \lim_{n \rightarrow \infty} \int_0^t \mathbb{E}^{\mathbb{Q}^d} [D_u \delta_n^a(X_u) Y_u^2] du \\ &= \lim_{n \rightarrow \infty} \int_0^t \mathbb{E}^{\mathbb{Q}^d} [\delta_n^a(X_u) \mathbb{E}^{\mathbb{Q}^d} [D_u Y_u^2 | X_u]] du. \end{aligned}$$

We denote $\gamma(x, u) = \mathbb{E}^{\mathbb{Q}^d} [D_u Y_u^2 | X_u = x]$, such that

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}^d} \left[\int_0^t D_u dl_u^a \right] &= \lim_{n \rightarrow \infty} \int_0^t \int_0^\infty \delta_n^a(x) \gamma(x, u) \phi(x, u) dx du \\ &= \lim_{n \rightarrow \infty} \int_0^\infty \delta_n^a(x) \left(\int_0^t \gamma(x, u) \phi(x, u) du \right) dx, \end{aligned}$$

where we have used Fubini's Theorem in the second line. By the continuity assumptions on ϕ and γ , we deduce that

$$\mathbb{E}^{\mathbb{Q}^d} \left[\int_0^t D_u dl_u^a \right] = \int_0^t \left(\mathbb{E}^{\mathbb{Q}^d} [D_u Y_u^2 | X_u = a] \phi(a, u) \right) du.$$

□

A.1.2 Proof of Lemma 3

Proof of Lemma 3. Since α and $\mathbf{1}_{S_u \geq K}$ are bounded, the process

$$M_t = \int_0^t \mathbf{1}_{S_u \geq K} D_u^d \alpha(S_u, u) S_u \sqrt{V_u} dW_u,$$

is a true martingale if

$$\mathbb{E}^{\mathbb{Q}^d} \left[\int_0^t (D_u^d S_u)^2 V_u du \right] < \infty.$$

On the one hand, since $t < T^*$, from Proposition 3.13 in [14], we can find $\omega > 2$ such that

$$\sup_{u \in [0, t]} \mathbb{E}^{\mathbb{Q}^d} \left[(D_u^d S_u)^\omega \right] < \infty.$$

On the other hand, from Theorem 3.1 in [28],

$$\sup_{u \in [0, t]} \mathbb{E}^{\mathbb{Q}^d} \left[V_u^{\frac{\omega}{\omega-2}} \right] < \infty.$$

Using Hölder's inequality with the pair $(\frac{\omega}{2}, \frac{\omega}{\omega-2})$,

$$\mathbb{E}^{\mathbb{Q}^d} \left[(D_u^d S_u)^2 V_u \right] \leq \mathbb{E}^{\mathbb{Q}^d} \left[(D_u^d S_u)^\omega \right]^{\frac{2}{\omega}} \mathbb{E}^{\mathbb{Q}^d} \left[V_u^{\frac{\omega}{\omega-2}} \right]^{\frac{\omega-2}{\omega}} < \infty.$$

Finally, using the Fubini theorem, $\mathbb{E}^{\mathbb{Q}^d} \left[\int_0^t (D_u^d S_u)^2 V_u du \right] < \infty$ and hence M is a true martingale of zero expectation. □

A.2 Local volatility calibration algorithm

A.2.1 Computation of the target volatility surface

For the calibration routine, we will compute the solution of the PDE (4.14) by a finite difference method. The spot grid is defined on $[0, S_{max}]$, where $S_{max} = S_0 e^{\frac{\sigma}{2} \sigma_{ATM}(\frac{T_{max}}{2}) \sqrt{\frac{T_{max}}{2}}}$. In order to speed up the calibration routine, we prefer not to use too many spot steps and time steps (150 steps in space and 20 time steps per year). Hence, the scheme will not have converged to the solution of the PDE at this point. In order to tackle this problem and still benefit from a good speed-up, we will compute a “target volatility surface”: instead of calibrating the market volatility surface, we will calibrate a volatility surface that takes into account the discretisation error of the numerical PDE solution. Industry practitioners like Murex use this approach [32]. The algorithm to build the target surface is explained below.

Algorithm 3 Computation of the target volatility surface

```

for (  $i = 1; i \leq N_{Mat}; i++$  ) do
  for (  $j = 1; j \leq M_i; j++$  ) do
    define  $\sigma_{Market} = \Sigma_{Market}(K_{i,j}, T_i)$  from the market volatility surface
    solve the PDE (4.14) with constant local vol  $\sigma_{LV} = \sigma_{Market}$ 
    get  $C(K_{i,j}, T_i)$  from the numerical solution
    get  $\Sigma_{Target}(K_{i,j}, T_i)$  by inverting the price with the Black-Scholes formula
  end for
end for

```

A.2.2 Calibration by fixed-point algorithm and forward induction

The calibration can be performed with a regularised minimisation using a penalised function. We refer the reader to [15, 32] for this approach. In this work, we choose a fixed-point algorithm for time efficiency, but want to highlight that a penalised minimisation yields to the same calibrated function σ_{LV} . The local volatility function is defined on a grid of points interpolated with cubic splines in spot and backward flat in time. In the FX case, where there are 5 quoted strikes per maturity (10 maturities), the local volatility is defined on a grid of 50 points. Each one of the points $\sigma_{LV}^{i,j} = \sigma_{LV}(K_{T_i,j}, T_i)$, with $i \in \llbracket 1, 10 \rrbracket$ and $j \in \llbracket 1, 5 \rrbracket$, can be seen as a parameter of the local volatility surface. For a given maturity T_i , the local volatility is defined on the interval $[K_{T_i,1}, K_{T_i,5}]$ and is extrapolated flat outside those bounds.

In order to define a first guess for the calibration routine, we use a smoothed bi-variate cubic spline following the algorithm in [17] to interpolate in strike and maturity the call prices on the market. This allows us to use the Dupire formula to define a first guess for the first maturity $T = T_1$. After the calibration of the first maturity pillar T_1 , the first guess for the next pillar is the current maturity local volatility. This approach has shown the best stability and speed in our tests.

As we now have a way to get the model implied volatility from the local volatility (with Φ), one can follow a Picard fixed-point algorithm as proposed in [37, 45] that we describe below.

Algorithm 4 Fixed-point forward induction

```

for (  $i = 1; i \leq N_{Mat}; i++$  ) do
  while  $it < \text{maxIter}$  do
    solve PDE (4.14) on  $[T_{i-1}, T_i]$ 
    compute model implied vol  $\Sigma_{Model}$  for maturity  $T_i$  from the computed call prices
    compute  $error = \sum_{m=1}^{M_i} (\Sigma_{Model}(K_{T_i,m}, T_i) - \Sigma_{Target}(K_{T_i,m}, T_i))^2$ 
    if  $error < \text{tol}$  then
      endwhile
    else
      for (  $j = 1; j \leq M_i; j++$  ) do
        update local volatility guess
        
$$\sigma_{LV}(K_{T_i,j}, T_i) = \sigma_{LV}(K_{T_i,j}, T_i) \frac{\Sigma_{Target}(K_{T_i,j}, T_i)}{\Sigma_{Model}(K_{T_i,j}, T_i)}$$

      end for
    end if
     $it++$ 
  end while
end for

```

Remark. As stated in [37], the map Φ is contracting and so is $f(\{\sigma_{LV}\}) \rightarrow \{\sigma_{LV}\} * \frac{\{\Sigma_{Target}\}}{\Phi(\{\sigma_{LV}\})}$. Then f admits a unique fixed point that is the limit of the sequence of local volatility guesses $(\sigma_{LV}^n)_{n \in \mathbb{N}}$ defined as $\{\sigma_{LV}^{n+1}\} = f(\{\sigma_{LV}^n\})$. Convergence is achieved particularly fast (between 10 and 20 iterations in practice).

The calibrated local volatility is shown in Figure A.1, where we plot it on a time scale to T_{max} for a better illustration of its shape.

We push the discretisation settings for the calibration to 800 space steps and 100 time steps per year (we use finite element method with quadratic basis functions to solve the PDE) and price quoted vanilla contracts with the backward Feynman-Kac PDE under the calibrated local volatility model. Then, we get a maximum error in implied volatility smaller than 0.01% (i.e., for a market volatility of 20%, the calibrated volatility could be $20.00 \pm 0.01\%$ in the worst case scenario).

Additionally, we plot the discounted marginal density of the spot extracted from the market. As mentioned before, this quantity is $\frac{\partial^2 C_{LV}}{\partial K^2}$ and can be computed from the PDE solution immediately and accurately. As we will use the density in the calibration formula in Theorem 3.1, we want it to be smooth and accurate. Figure A.2 shows that quantity.

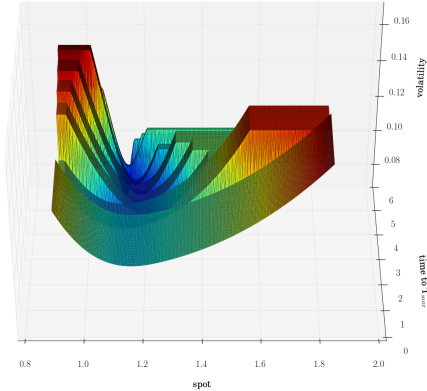


Figure A.1: EURUSD Local volatility function calibrated by forward PDE and fixed-point algorithm

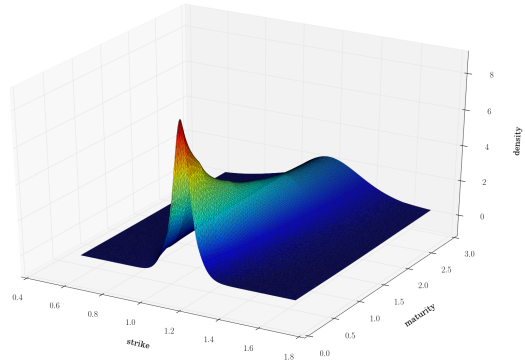


Figure A.2: Market spot marginal density computed from the Dupire forward PDE with calibrated σ_{LV}

A.3 Shifted CIR model calibration

In order to estimate the zero coupon curve (also known as the term structure of interest rates or the yield curve), we assume that the instantaneous forward rate is piecewise-flat. Consider the time nodes $t_0=0, t_1, \dots, t_n$ and the set of estimated instantaneous forward rates f_1, f_2, \dots, f_n from which the curve is constructed, and define

$$\bar{r}(t) = f_i \quad \text{if} \quad t_{i-1} \leq t < t_i, \quad \text{for} \quad i = 1, 2, \dots, n. \quad (\text{A.5})$$

Using (4.9) – (A.5) and solving the resulting linear system of equations, we get

$$f_i = \frac{1}{\Delta(t_{i-1}, t_i)} \ln \left(\frac{1 + \Delta(0, t_i)R(0, t_i)}{1 + \Delta(0, t_{i-1})R(0, t_{i-1})} \right) \quad \text{for} \quad i = 1, 2, \dots, n. \quad (\text{A.6})$$

The continuously-compounded spot rate, i.e., the constant rate at which the value of a pure discount bond must grow to yield one unit of currency at maturity, is defined as

$$R_0(0, t) = \frac{1}{\Delta(0, t)} \int_0^t \bar{r}(s) ds. \quad (\text{A.7})$$

Using (A.5) – (A.7), we deduce that

$$R_0(0, t) = \frac{\Delta(t, t_i)}{\Delta(0, t)\Delta(t_{i-1}, t_i)} \ln \left(1 + \Delta(0, t_{i-1})R(0, t_{i-1}) \right) + \frac{\Delta(t_{i-1}, t)}{\Delta(0, t)\Delta(t_{i-1}, t_i)} \ln \left(1 + \Delta(0, t_i)R(0, t_i) \right) \quad (\text{A.8})$$

whenever $t_{i-1} \leq t < t_i$. In Figure A.3, we plot the USD and EUR zero coupon curves $t \mapsto R_0(0, t)$, $t > 0$, estimated from the quoted deposit rates from March 18, 2016, together with the flat-forward instantaneous forward rates.

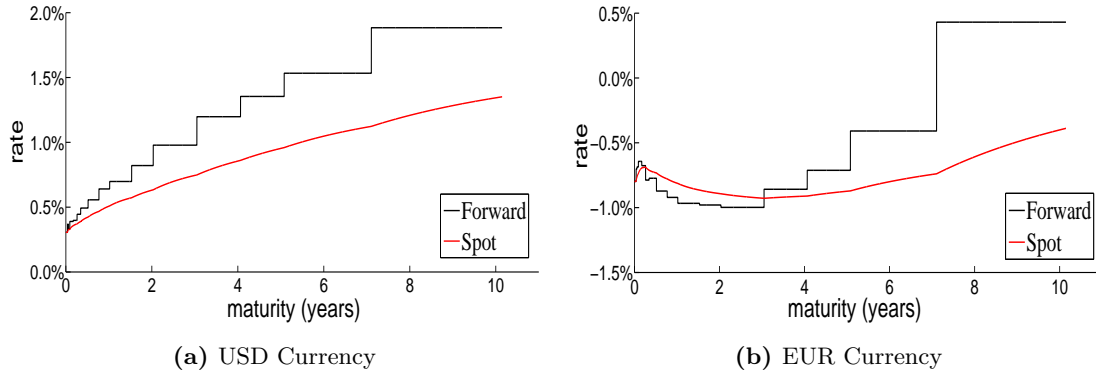


Figure A.3: The instantaneous forward rates and the continuously-compounded spot rates.

A choice of the shift function h as in (4.8) results in an exact fit to the initial term structure of interest rates independent of the value of the parameter vector β_1 .

Next, we determine β_1 by calibrating the CIR++ model to the current term structure of volatilities, in particular, by fitting at-the-money (ATM) cap volatilities. We consider caps with integer maturities ranging from 1 to 10 years for both currencies, with an additional 18 month cap for EUR. For USD, all caps have quarterly frequency, whereas for EUR the 1 year and 18 month caps have quarterly frequency and the 2 to 10 year caps have semi-annual frequency. A cap is a set of spanning caplets with a common strike so the value of the cap is simply the sum of the values of its caplets. It is market standard to price caplets with the Black formula, in which case the fair value of the cap at time 0 with rate (strike) K , reset times $T_a, T_{a+1}, \dots, T_{b-1}$ and payment times $T_{a+1}, \dots, T_{b-1}, T_b$ is:

$$\text{Cap}_{\text{Black}}(K, \sigma_{a,b}) = \sum_{i=a+1}^b P(0, T_i) \Delta(T_{i-1}, T_i) \text{Black}(K, F(0, T_{i-1}, T_i), \sigma_{a,b} \sqrt{T_{i-1}}), \quad (\text{A.9})$$

where $F(0, T, S)$ is the simply-compounded forward rate at time 0 for the expiry T and maturity S defined as

$$F(0, T, S) = \frac{1}{\Delta(T, S)} \left(\frac{P(0, T)}{P(0, S)} - 1 \right) \quad (\text{A.10})$$

and the Black volatility $\sigma_{a,b}$ corresponding to a strike K is retrieved from market quotes. Denoting by ϕ_0 and Φ_0 the standard normal probability density function (PDF) and cumulative distribution function (CDF), respectively, Black's formula is:

$$\begin{aligned} \text{Black}(K, F, v) &= F\Phi_0(d_1) - K\Phi_0(d_2), \\ d_{1,2} &= \frac{\ln(F/K) \pm v^2/2}{v}. \end{aligned} \quad (\text{A.11})$$

However, Black's formula cannot cope with negative forward rates F or strikes K , in which case we switch to Bachelier's (normal) formula in (A.9):

$$\begin{aligned} \text{Normal}(K, F, v) &= (F - K)\Phi_0(d) + v\phi_0(d), \\ d &= \frac{F - K}{v}. \end{aligned} \quad (\text{A.12})$$

The data in Figure A.3b suggest that the instantaneous forward rate for EUR takes negative values. Therefore, we use Black cap volatility quotes for USD and Normal cap volatility quotes for EUR. The market prices of at-the-money caps are computed by inserting the forward swap rate

$$S_{a,b} = \frac{P(0, T_a) - P(0, T_b)}{\sum_{i=a+1}^b \Delta(T_{i-1}, T_i) P(0, T_i)} \quad (\text{A.13})$$

as strike and the quoted cap volatility as $\sigma_{a,b}$ in (A.9), using either Black's or Bachelier's formula.

Fitting the CIR++ model to cap volatilities means finding the value of β_1 for which the model cap prices, which are available in closed-form [6], best match the market cap prices. The calibration is performed by minimising the sum of the squared differences between model- and market-implied cap volatilities:

$$\min_{\beta_1 \in \mathbb{R}_+^4} \sum_{1 \leq i \leq n} [\sigma^{\text{CIR}}(T_i; \beta_1) - \sigma^{\text{M}}(T_i)]^2, \quad (\text{A.14})$$

where σ^{CIR} and σ^{M} stand for the model- and the market-implied cap volatilities, respectively, and T_1, \dots, T_n are the cap maturities. Model-implied cap volatilities are obtained by pricing market caps with the CIR++ model and then inverting the formula (A.9) in order to retrieve the implied volatility associated with each maturity. We choose to calibrate the model to cap volatilities since they are of similar magnitude, unlike cap prices which can differ by a few orders of magnitude. The calibration results are displayed in Table 1.

On the one hand, (A.14) is a highly nonlinear and non-convex optimisation problem, and the objective function may have multiple local minima. On the other hand, global optimisation algorithms require a very high computation time and do not scale well with complexity, as opposed to local optimisation methods. A fast calibration is important in practice since option pricing models may need to be re-calibrated several times within a short time span. Therefore, we used a nonlinear least-squares solver, in particular the trust-region-reflective algorithm [12], for the calibration and a global optimisation method, in particular a genetic algorithm [42], for verification purposes only.

Figure A.4 shows the fitting capability of the CIR++ model, and the implied cap volatility curve is compared to the market curve for each currency. Taking into account that the model has only 4 parameters to fit between 10 and 11 data points, we conclude that the CIR++ model provides a fairly reasonable fit to the term structure of cap volatilities $T_i \mapsto \sigma^{\text{M}}(T_i)$, $1 \leq i \leq n$.

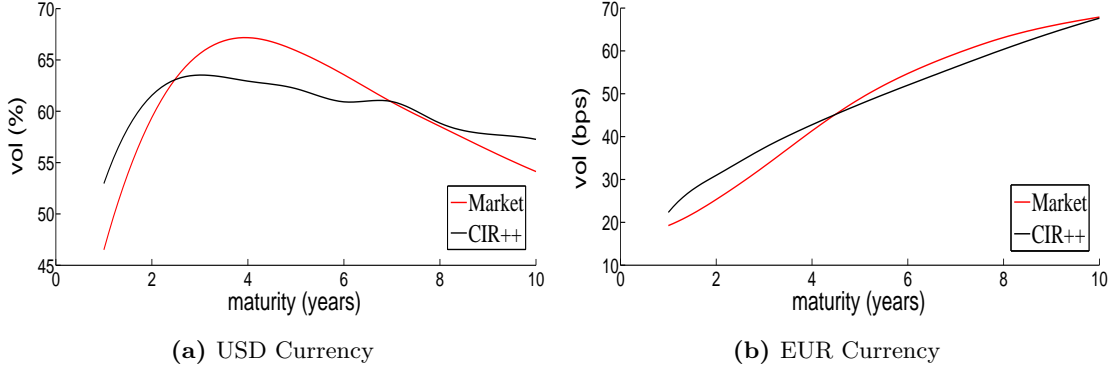


Figure A.4: The market- and model-implied term structures of cap volatilities.

A.4 Monte Carlo QE-scheme

The *Quadratic-Exponential (QE)* scheme [3] uses moment-matching techniques and can significantly reduce the Monte Carlo discretisation error. While the full truncation Euler scheme and the QE scheme have shown to perform well in our tests, we experienced a faster convergence in time for the QE scheme when the Feller condition is broken. Hence, we choose the QE scheme for the variance process and the full truncation Euler for both stochastic rates, as the computational cost will be smaller. We briefly write a generalised QE scheme based on the original scheme from [3] to incorporate a leverage function and stochastic rates in the discretisation.

We follow our time interpolation rule for the calibration of α and interpolate forward-flat in time. We assume for simplicity that each Monte Carlo time step belongs to the α time grid. We can write

$$V_{t+\Delta t} = V_t + \int_t^{t+\Delta t} \kappa (\theta - V_u) du + \xi \int_t^{t+\Delta t} \sqrt{V_u} dW_u^V,$$

and hence

$$\int_t^{t+\Delta t} \sqrt{V_u} dW_u^V = \frac{V_{t+\Delta t} - V_t - \int_t^{t+\Delta t} \kappa (\theta - V_u) du}{\xi},$$

and

$$d \ln S_t = \left(r_t^d - r_t^f - \frac{1}{2} \alpha^2 (S_t, t) V_t \right) dt + \alpha (S_t, t) \rho \sqrt{V_t} dW_t^V + \alpha (S_t, t) \sqrt{1 - \rho^2} \sqrt{V_t} dW_t^S,$$

where W_t^S is a Brownian motion independent of W_t^V . Therefore,

$$\begin{aligned} \ln S_{t+\Delta t} &= \ln S_t + \int_t^{t+\Delta t} (r_u^d - r_u^f) du - \frac{1}{2} \alpha^2 (S_t, t) \int_t^{t+\Delta t} V_u du \\ &\quad + \frac{\alpha (S_t, t) \rho \left(V_{t+\Delta t} - V_t - \kappa \theta \Delta t + \kappa \int_t^{t+\Delta t} V_u du \right)}{\xi} \\ &\quad + \alpha (S_t, t) \sqrt{1 - \rho^2} \int_t^{t+\Delta t} \sqrt{V_u} dW_u^S. \end{aligned}$$

We approximate $\int_t^{t+\Delta t} V_u du$ by $\left(\frac{V_{t+\Delta t} + V_t}{2} \right) \Delta t$, and note that conditional on V_t and $\int_t^{t+\Delta t} V_u du$, since W_u^V and W_u^S are independent, the Itô integral $\int_t^{t+\Delta t} \sqrt{V_u} dW_u^S$ is normally distributed with mean zero and variance $\int_t^{t+\Delta t} V_u du$. We write the full scheme below

$$\left\{ \begin{array}{l} g_{t+\Delta t}^d = g_t^d + \kappa_d \left(\theta_d - (g_t^d)^+ \right) \Delta t + \xi_d \sqrt{(g_t^d)^+} \sqrt{\Delta t} Z_d \\ g_{t+\Delta t}^f = g_t^f + \kappa_f \left(\theta_f - (g_t^f)^+ \right) \Delta t + \xi_f \sqrt{(g_t^f)^+} \sqrt{\Delta t} Z_f \\ \ln S_{t+\Delta t} = \ln S_t + \left(\frac{(r_{t+\Delta t}^d - r_{t+\Delta t}^f) + (r_t^d - r_t^f)}{2} - \frac{1}{4} \alpha^2(S_t, t) (V_{t+\Delta t} + V_t) \right) \Delta t \\ \quad + \frac{\alpha(S_t, t) \rho (V_{t+\Delta t} - V_t + \kappa \left(\frac{V_{t+\Delta t} + V_t}{2} - \theta \right) \Delta t)}{\xi} \\ \quad + \alpha(S_t, t) \sqrt{1 - \rho^2} \sqrt{\frac{V_{t+\Delta t} + V_t}{2}} \sqrt{\Delta t} Z \\ V_{t+\Delta t} : \quad \begin{cases} \text{if } \psi \leq \psi_c & : V_{t+\Delta t} = a (b + Z_v)^2 \\ \text{else} & \begin{cases} \text{if } U \leq p & : V_{t+\Delta t} = 0 \\ \text{else} & : V_{t+\Delta t} = \ln \left(\frac{1-p}{1-U} \right) \frac{m}{1-p}, \end{cases} \end{cases} \end{array} \right. \quad (\text{A.15})$$

with

$$\left\{ \begin{array}{l} m = \theta + (V_t - \theta) e^{-\kappa \Delta t}, \\ \gamma^2 = \frac{V_t \xi^2 e^{-\kappa \Delta t}}{\kappa} (1 - e^{-\kappa \Delta t}) + \frac{\theta \xi^2}{2\kappa} (1 - e^{-\kappa \Delta t})^2, \\ \psi = \frac{\gamma^2}{m^2}, \quad p = \frac{\psi-1}{\psi+1}, \quad \beta = \frac{1-p}{m}, \\ b^2 = \frac{2}{\psi} - 1 + \sqrt{\frac{2}{\psi} \sqrt{\frac{2}{\psi} - 1}}, \quad a = \frac{m}{1+b^2}, \\ \psi_c = 1.5, \end{array} \right.$$

where Z, Z_v, Z_d, Z_f are independent draws from a standard normal distribution and U is a draw from a uniform distribution.

A.5 Finite element mesh construction

For an accurate finite element solution, the mesh is of paramount importance. In order to refine the mesh in the most relevant area, we use an exponential mesh on the variance axis and a hyperbolic mesh (see [50]) in the spot direction. This makes the mesh finer around $z = 0$ and $x = S_0$. In order to build our mesh, we first define the grids in spot $(x_i)_{i \in \llbracket 0, N_S \rrbracket}$ and variance $(z_j)_{j \in \llbracket 0, N_V \rrbracket}$ separately. Additionally, to solve the PDE numerically, we need to truncate at the boundary and use $\bar{\Omega} = \{(x, z) \in [0, S_{max}] \times [0, V_{max}]\}$ on a time interval $[0, T]$. We choose

$$S_{max} = S_0 e^{5\alpha(S_0, T) \sqrt{(v_0 e^{-\kappa T} + \theta(1 - e^{-\kappa T}))T}}$$

and recall that the stationary distribution of the CIR process is a gamma distribution of density ϕ_v^∞ as defined in (4.17). We compute V_{max} with the inverse cumulative density function such that

$$\mathbb{P}(z > V_{max}) = 0.01\%.$$

We write

$$\begin{aligned} x_i &= f_h(g_h(\bar{x}_i)), \\ z_j &= f_e(g_e(\bar{z}_j)), \end{aligned}$$

with

$$\begin{aligned}
f_h(x) &= S_0 + b \sinh(\nu x + d), \\
b &= \eta(S_{max} - S_{min}), \\
d &= \operatorname{arcsinh}\left(\frac{S_{min} - S_0}{b}\right), \\
\nu &= \operatorname{arcsinh}\left(\frac{S_{max} - S_0}{b}\right) - d, \\
\bar{x}_i &= \frac{i}{(N_S + 1)}, \\
\eta &= 0.02,
\end{aligned}
\qquad
\begin{aligned}
f_e(z) &= c + c \exp(\lambda z), \\
c &= \frac{V_{max}}{e^\lambda - 1}, \\
\lambda &= \max\left(1, 4 - \frac{3\kappa\theta}{\xi^2}\right), \\
\bar{z}_j &= \frac{j}{(N_V + 1)},
\end{aligned}$$

where η is defined according to our numerical experiments and g_h is the quadratic polynomial that passes through the points $(0, 0)$, $(1, 1)$,

where λ is defined according to our numerical experiments and g_h is the quadratic polynomial that passes through the points $(0, 0)$, $(1, 1)$,

$$\left(\frac{\lfloor f_h^{-1}(S_0)(N_S + 1) + 0.5 \rfloor}{N_S + 1}, f_h^{-1}(S_0) \right). \qquad \left(\frac{\lfloor f_e^{-1}(v_0)(N_V + 1) + 0.5 \rfloor}{N_V + 1}, f_e^{-1}(v_0) \right).$$

The latter intermediate step makes sure that both S_0 and v_0 are vertices of their respective grids. The construction of the finite element triangular mesh can be achieved by creating a vertex at each point (x_i, z_j) and defining two triangular cells (upper left and lower right) in each rectangle. The finite element solution is solved with \mathbb{P}_2 elements, i.e., a polynomial of order two on a triangle cell. Each triangle is characterised by 6 local degrees of freedom (nodes) as displayed in Figure A.5 (see [33] for details). An example of a generated mesh with 30 spot steps and 30 variance steps is illustrated in Figure A.6.

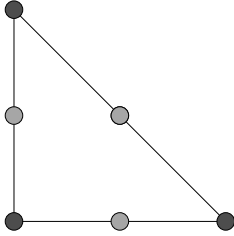


Figure A.5: \mathbb{P}_2 element with 6 degrees of freedom

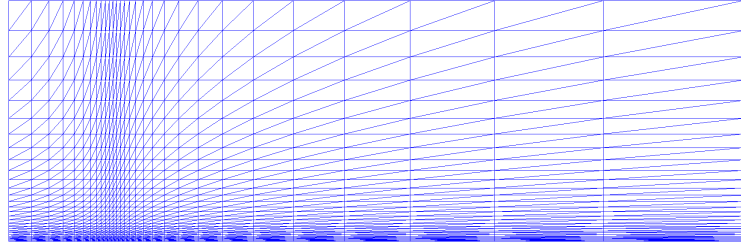


Figure A.6: Finite element triangular mesh refined around $x = S_0$ and $z = 0$

In order to see the improvement due to the changed-of-variable PDE (4.18) for p , we plot both $v^{-\beta}\phi$ and p for a pure Heston model with

$$r = 3\%, \quad q = 1\%, \quad \kappa = 1, \quad \theta = v_0 = 0.04, \quad \rho = -0.3, \quad \xi = 0.5, \quad t = 1,$$

which corresponds to a Feller ratio of 0.32. We use 100 time steps as well as 30 spot steps and 30 variance steps. The solution for $p = v^{-\beta}\phi$ computed with no change of variables is presented in Figure A.7 where we notice significant numerical instabilities. The change of variables PDE is solved for the same problem and p is plotted in Figure A.8, where the improvement is clear and the solution is smooth and accurate.

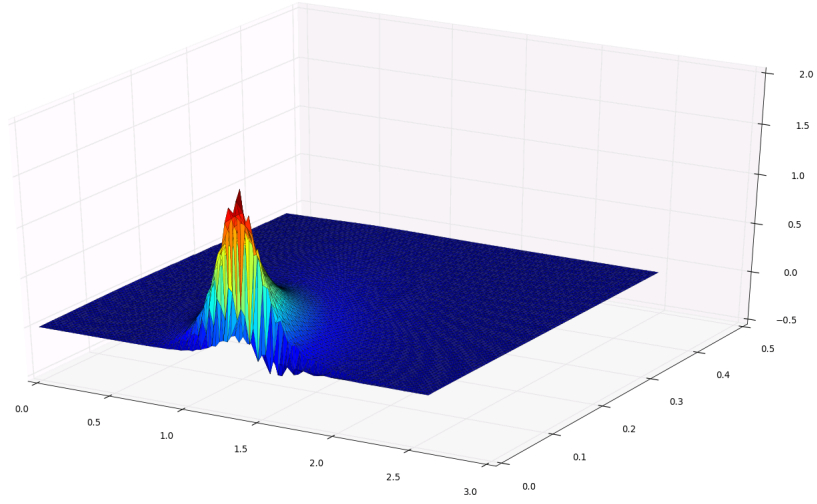


Figure A.7: $p = v^{-\beta} \phi$ using ϕ computed with the standard PDE

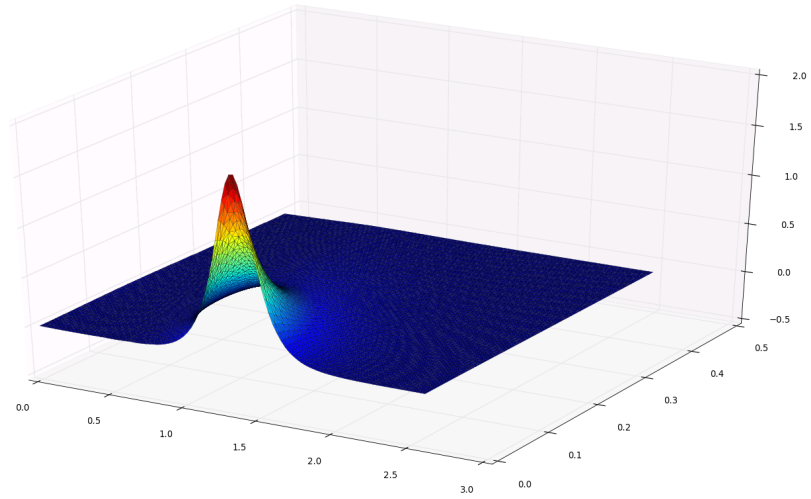


Figure A.8: $p = v^{-\beta} \phi$ computed directly from the change of variables PDE